

全球大気SPEEDYモデルの実習 (SPEEDY-LETKF)

小槻 峻司

理化学研究所 計算科学研究機構



With many thanks to
RIKEN Data Assimilation Research Team
(特に, 近藤さんと岡崎さん!!!)



iTHES データ同化スクール, Sep. 12, 2016 @ 理研IIB (神戸)

資料1

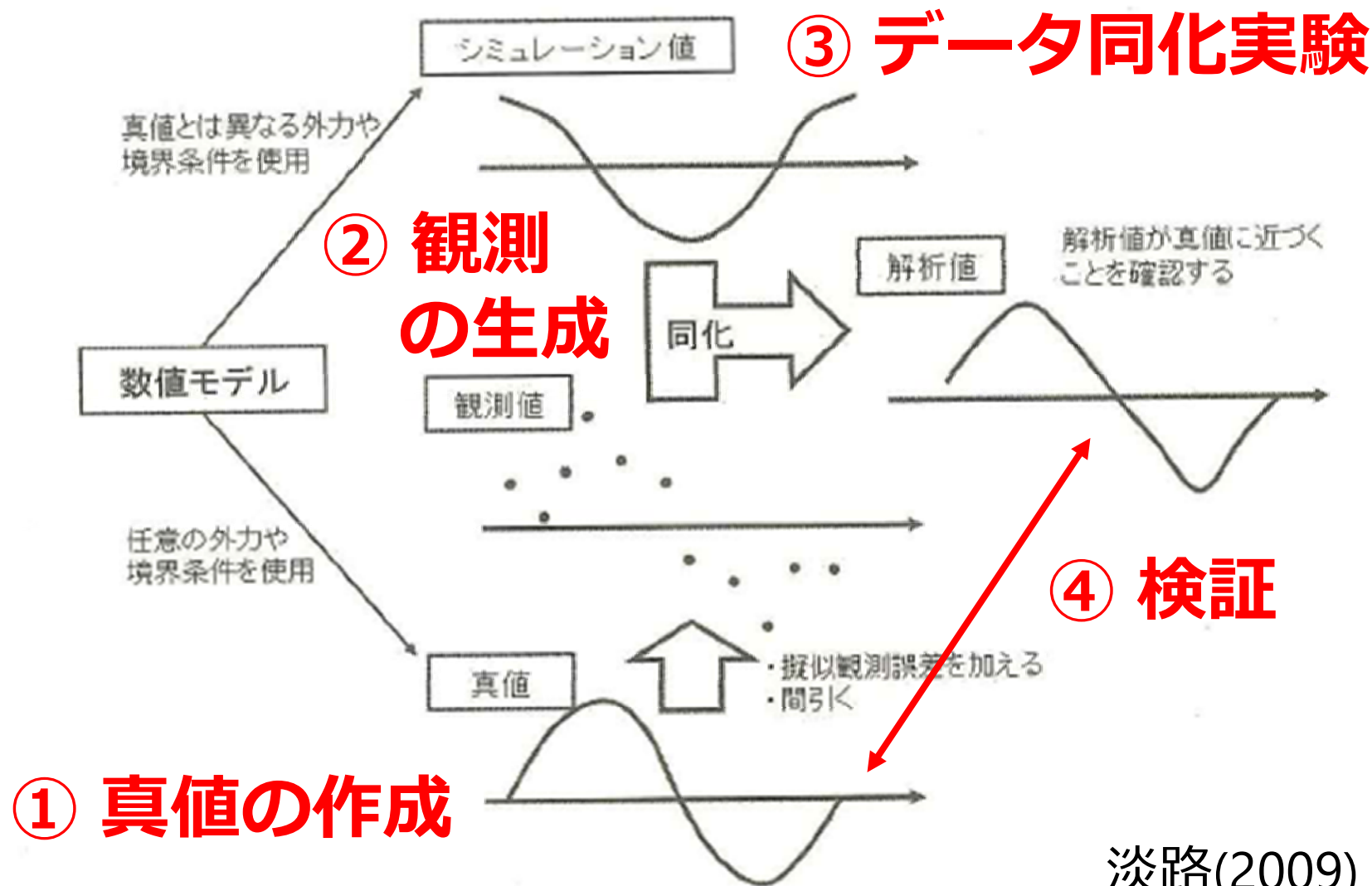
- スケジュール
- 配布資料の確認
- 注意事項
- 実習班と課題

実習の目標

- OSSEを理解する
- データ同化の基本を理解する
- 重要なパラメータの役割を理解する
- SPEEDY-LETKFを自分で回せるようになる

OSSE (観測システムシミュレーション実験)

別に，双子実験 (idealized twin experiment)とも呼ばれる

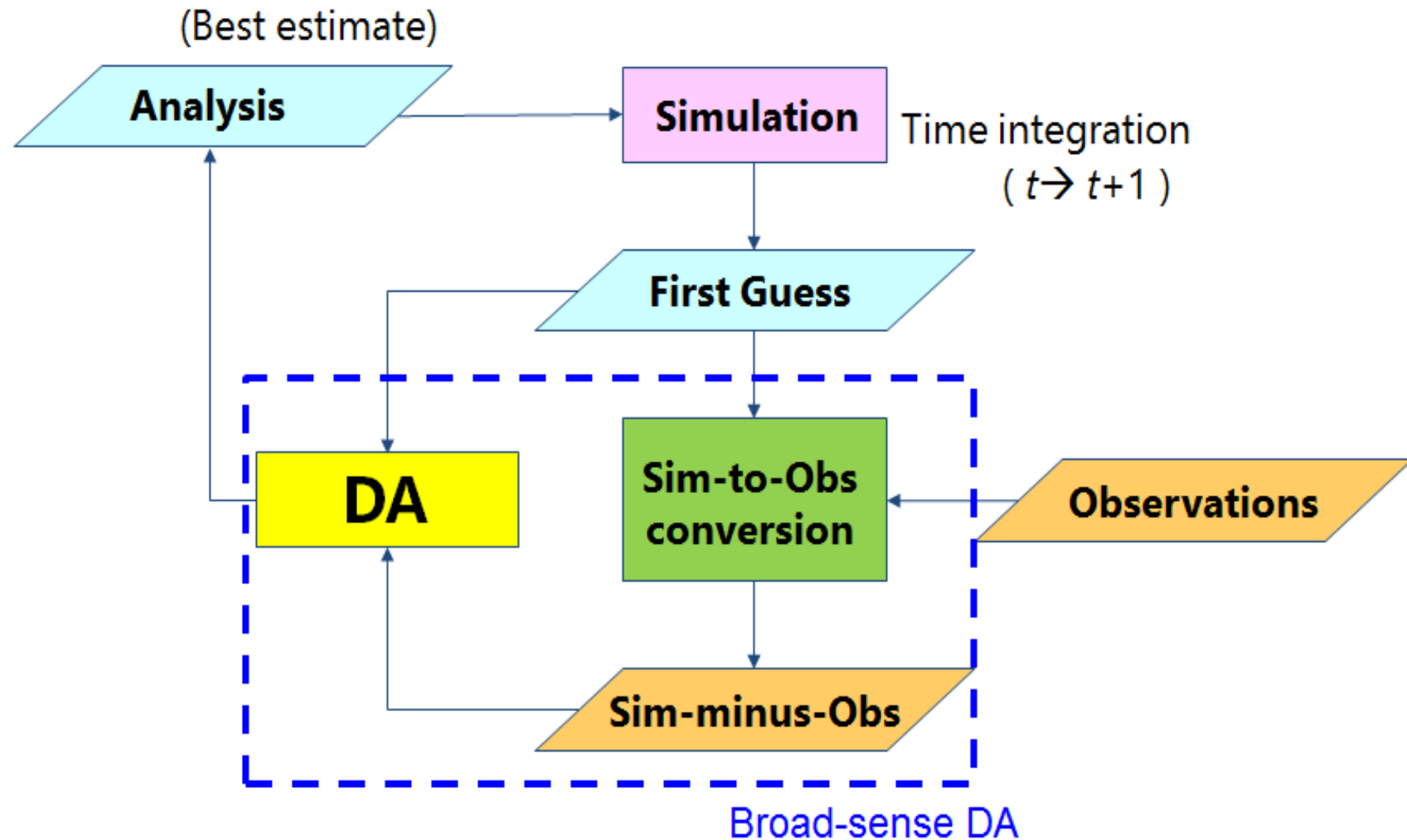


OSSEのメリット

- 真値を知っている
- 観測誤差が既知である
- モデルが完全である
- 架空の状況を設定可能

データ同化のコンセプト

データ同化サイクル



誤差共分散行列

平均
(mean)

$$E[\mathbf{x}] = \sum_{i=1}^n \mathbf{x}_i$$

$E[\]$: 統計的期待値

分散
(variance)

$$VAR[\mathbf{x}] = E\left[(\mathbf{x} - E[\mathbf{x}])^2\right]$$

共分散
(covariance)

$$COV[\mathbf{x}, \mathbf{y}] = E\left[(\mathbf{x} - E[\mathbf{x}])(\mathbf{y} - E[\mathbf{y}])\right]$$

相関
(correlation)

$$CORR[\mathbf{x}, \mathbf{y}] = COV[\mathbf{x}, \mathbf{y}] / \sqrt{VAR[\mathbf{x}] \cdot VAR[\mathbf{y}]}$$

誤差共分散行列

$$\delta \mathbf{x} = \mathbf{x} - \mathbf{x}^{truth}$$

$$\mathbf{P} = E\left[\delta \mathbf{x} [\delta \mathbf{x}]^T\right]$$

対角成分: variance
推定変数の確からしさ

非対角成分: covariance
変数間の関係

Kalman Filter (wo/ model error)

$$\mathbf{x}_t^f = \mathbf{M}\mathbf{x}_{t-1}^a$$

$$\mathbf{P}_t^f = \mathbf{M}\mathbf{P}_{t-1}^a\mathbf{M}^T$$

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

$$\mathbf{x}_t^a = \mathbf{x}_t^f + \mathbf{K}_t \left[\mathbf{y}_t^o - \mathbf{H}\mathbf{x}_t^f \right]$$

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

1. 時間発展 x: 状態変数
 M: モデル

2. 誤差の発展
 P: 誤差共分散

3. カルマンゲインKの取得
 H: 観測演算子
 R: 観測誤差共分散

4. 状態の解析

5. 誤差の解析

解析の式

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

x: 推定値

y: 観測値

P: 推定誤差共分散行列

R: 観測誤差共分散行列

H: 観測演算子

f: first guess

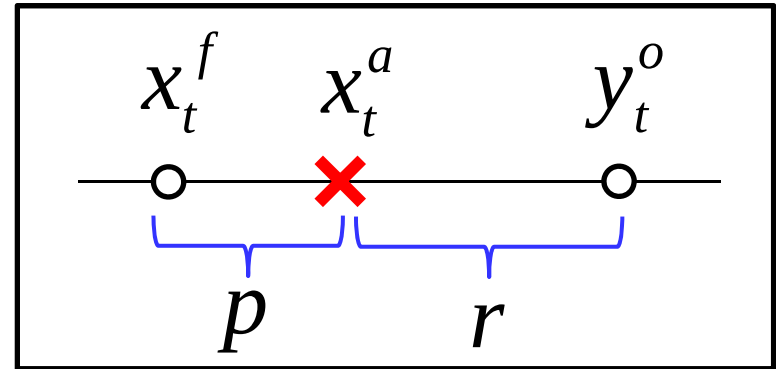
a: analysis

o: observation

簡単な例

もし一変数の問題だったら？ (Hは変数を直接観測と想定)

$$x_t^a = x_t^f + \frac{p}{p+r} [y_t^o - x_t^f]$$



結局は、行列で重み付け平均を計算している

➔ 大事なことは、 p と r を確からしく設定すること

実習課題

SPEEDY-LETKFによる
データ同化感度実験

SPEEDY-LETKF

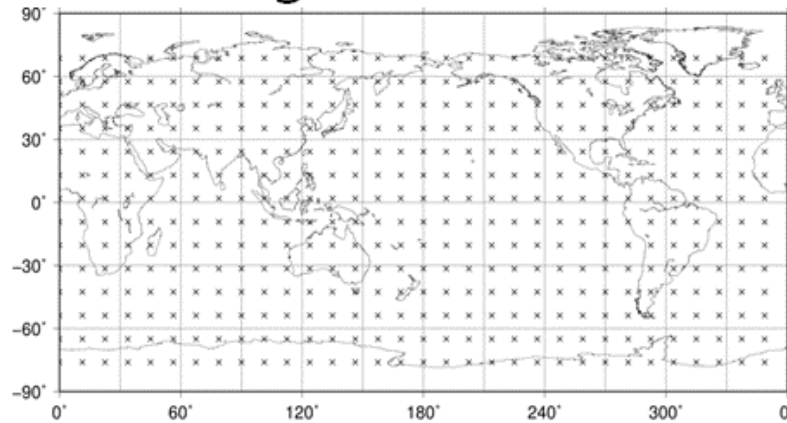
SPEEDY: 簡易な全球大気モデル

LETKF: EnKFのひとつであるデータ同化

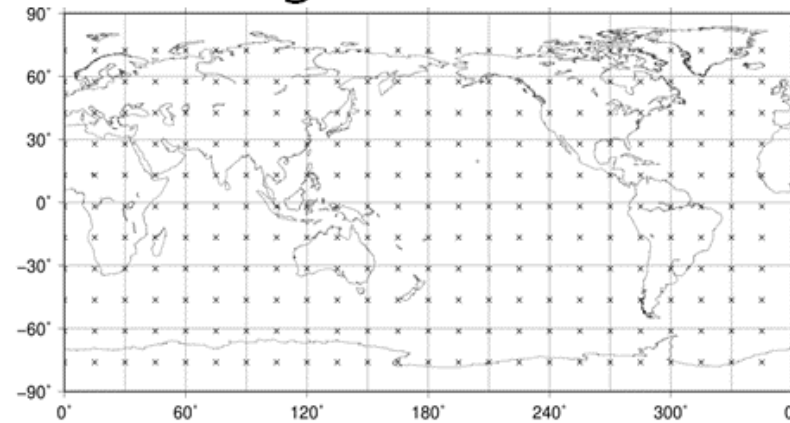
- 感度実験
 - 観測分布
 - 観測誤差
 - インフレーション
 - アンサンブルサイズ
 - 局所化スケール

観測分布を変えた感度実験

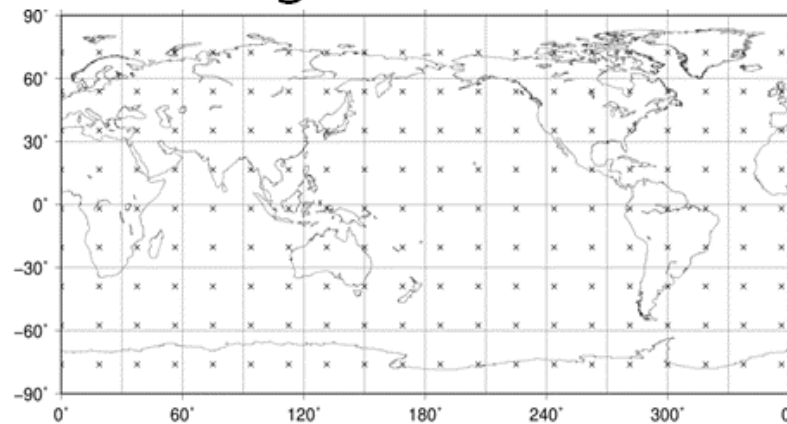
(a) reg3: # obs = 448



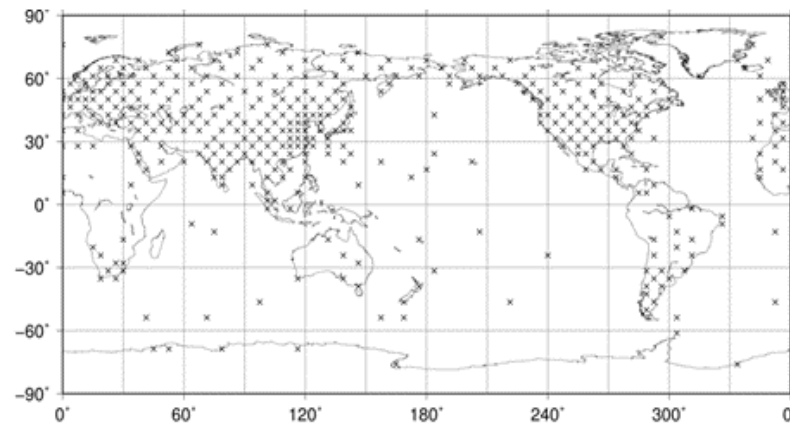
(b) reg4: # obs = 264



(c) reg5: # obs = 180



(d) raob: # obs = 416



実験する観測分布

観測誤差

OSSEでは, 自分で作成しているため**R**は既知.
実際の問題では, **R**は未知である.

$$\mathbf{R} \leftarrow \gamma \times \mathbf{R}^{truth}$$

γ : 感度実験のパラメータ

このパラメータを変えて実験する

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

$\mathbf{x}$: 推定値
 $\mathbf{y}$: 観測値

$\mathbf{P}$: 推定誤差共分散行列
 $\mathbf{R}$: 観測誤差共分散行列
 $\mathbf{H}$: 観測演算子

f : first guess
 a : analysis
 o : observation

インフレーション

EnKFでは、背景誤差共分散行列が過小評価される
(モデルの非線形性, アンサンブルメンバーの不足により)

$$\mathbf{P}_{inf}^f \leftarrow \Delta \times \mathbf{P}^f$$

Δ : inflation parameter
このパラメータを変えて実験する

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

$\mathbf{x}$: 推定値
 $\mathbf{y}$: 観測値

$\mathbf{P}$: 推定誤差共分散行列
 $\mathbf{R}$: 観測誤差共分散行列
 $\mathbf{H}$: 観測演算子

f : first guess
 a : analysis
 o : observation

アンサンブルサイズ

背景誤差共分散行列をアンサンブルメンバーで近似
→ スプレッドの大きさで確からしさを測る

$$\mathbf{X} = \left[\left(\mathbf{x}_1 - \bar{\mathbf{x}} \right), \left(\mathbf{x}_2 - \bar{\mathbf{x}} \right), \quad \dots, \left(\mathbf{x}_m - \bar{\mathbf{x}} \right) \right]$$

$$\mathbf{P} = \frac{1}{m-1} \mathbf{X} (\mathbf{X})^T$$

m : アンサンブルメンバー

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

$\mathbf{x}$: 推定値
 $\mathbf{y}$: 観測値

$\mathbf{P}$: 推定誤差共分散行列
 $\mathbf{R}$: 観測誤差共分散行列
 $\mathbf{H}$: 観測演算子

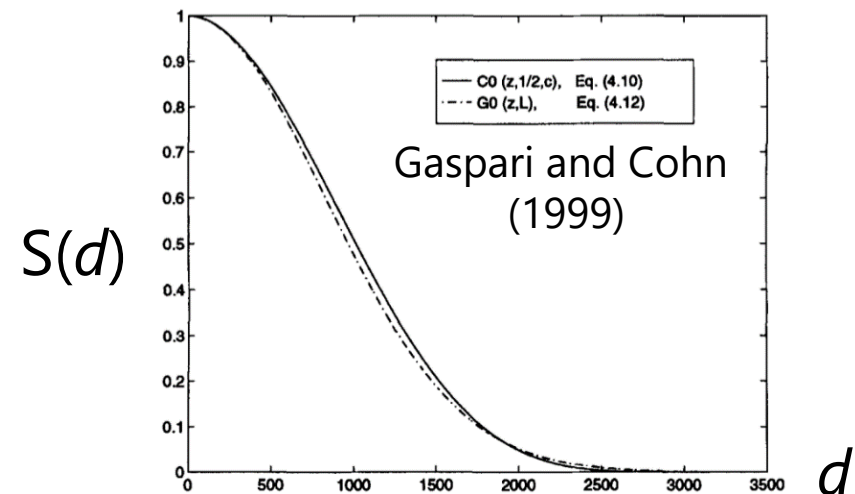
f : first guess
 a : analysis
 o : observation

局所化 (localization)

遠くの観測の情報が入りすぎないようにする
(サンプリングエラーの緩和)

$$\mathbf{P} = \frac{1}{m-1} \mathbf{X} (\mathbf{X})^T$$

$$\mathbf{P} \leftarrow S(d) \circ \mathbf{P}$$



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

$\mathbf{x}$: 推定値
 $\mathbf{y}$: 観測値

$\mathbf{P}$: 推定誤差共分散行列
 $\mathbf{R}$: 観測誤差共分散行列
 $\mathbf{H}$: 観測演算子

f : first guess
 a : analysis
 o : observation

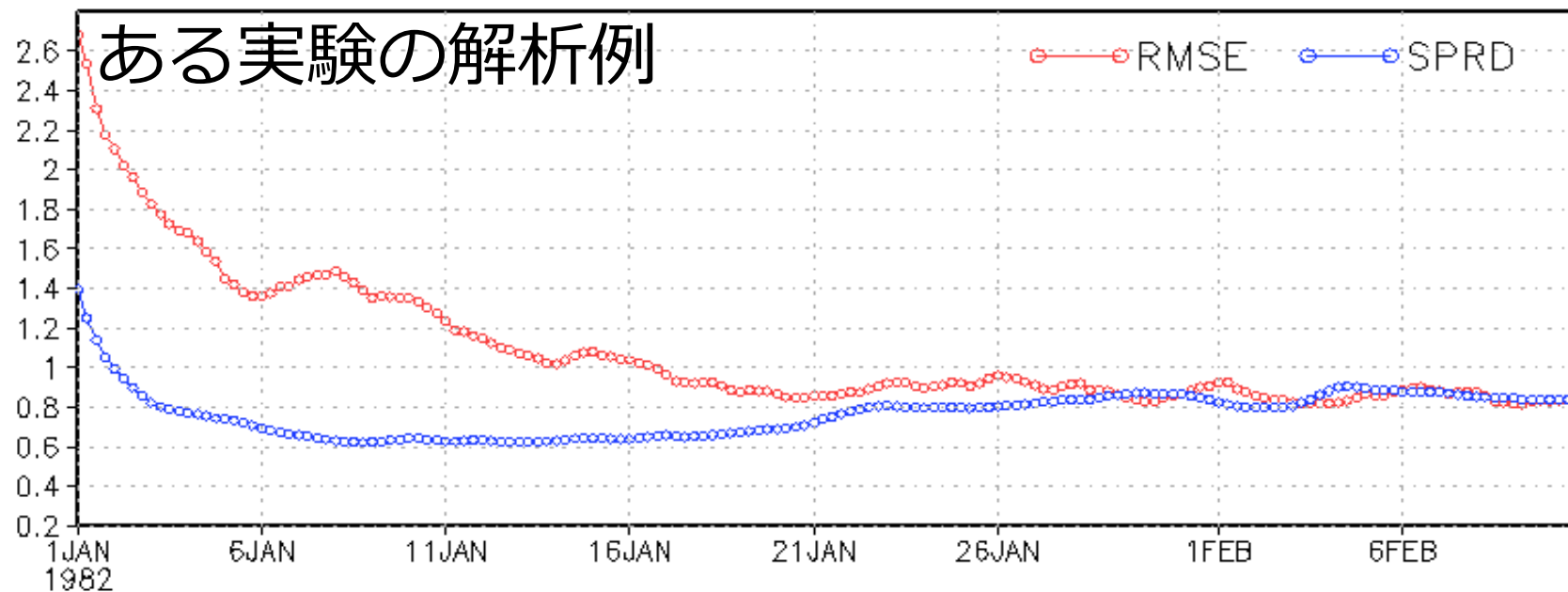
実習の流れ

- FX-10 (or 自前サーバー)にログイン
- SPEEDY-LETKFを一度実行 (OSSE)
- 感度実験
- 感度実験結果の解析

結果の解析

- 3ツールを準備
 - 時系列のRMS Error とアンサンブルスプレッド
 - RMS Errorの空間分布
 - Adaptive inflation の時系列

SPEEDY-LETKF : T [K] (4th model level ; 500 hPa)
From 1982010100 To 1982021018



まとめ

- OSSEを理解する
- データ同化の基本を理解する
- 重要なパラメータの役割を理解する
- SPEEDY-LETKFを自分で回せるようになる

Adaptive inflation methods

iTHES Data Assimilation School
(Sep 13, 2016)

Shunji KOTSUKI

Data Assimilation Research Team, RIKEN-AICS

Today's goals

- Review: covariance inflation methods
- Understand innovation statistics
- Understand adaptive inflation methods

Kalman Filter

$$\mathbf{x}_t^f = \mathbf{M}\mathbf{x}_{t-1}^a - \boldsymbol{\eta}$$

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

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

$$\mathbf{x}_t^a = \mathbf{x}_t^f + \mathbf{K}_t \left[\mathbf{y}_t^o - \mathbf{H}\mathbf{x}_t^f \right]$$

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

1. 時間発展

$\mathbf{x}$: 状態変数

$\mathbf{M}$: モデル

$\boldsymbol{\eta}$: ランダム誤差

2. 誤差の発展

$\mathbf{P}$: 誤差共分散

$\mathbf{Q}$: モデル誤差共分散

3. カルマンゲイン $\mathbf{K}$ の取得

$\mathbf{H}$: 観測演算子

$\mathbf{R}$: 観測誤差共分散

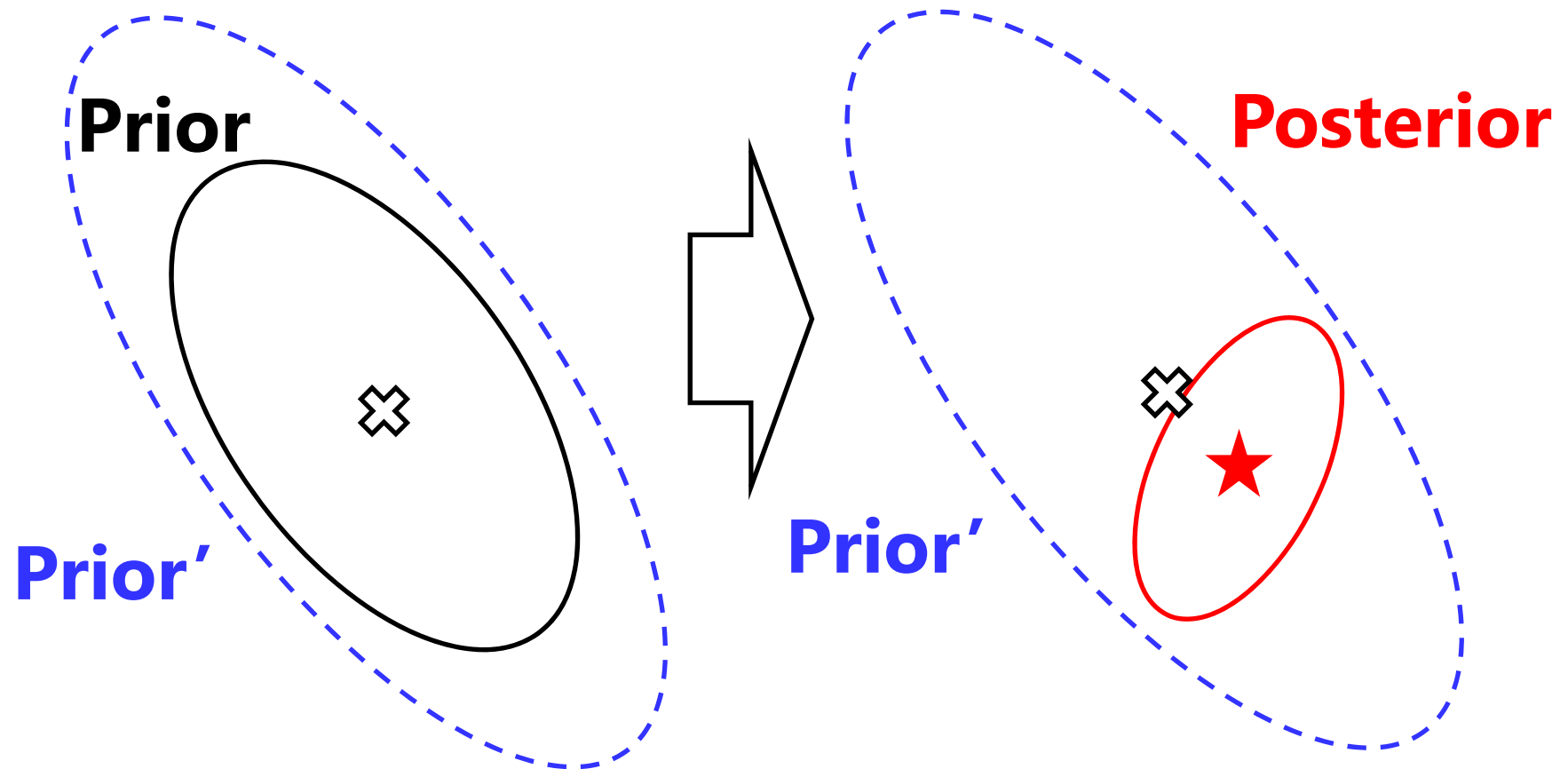
4. 状態の解析

5. 誤差の解析

Multiplicative inflation

Multiplicative inflation

$$\mathbf{P}^{b'} = \Delta \times \mathbf{P}^b$$



Covariance inflations in Kalman filters

$$\mathbf{x}_t^f = \mathbf{M}\mathbf{x}_{t-1}^a - \boldsymbol{\eta}$$

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

Multiplicative inflation

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

$$\mathbf{x}_t^a = \mathbf{x}_t^f + \mathbf{K}_t [\mathbf{y}_t^o - \mathbf{H}\mathbf{x}_t^f]$$

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

1. 時間発展

x: 状態変数

M: モデル

η : ランダム誤差

2. 誤差の発展

P: 誤差共分散

Q: モデル誤差共分散

3. カルマンゲインKの取得

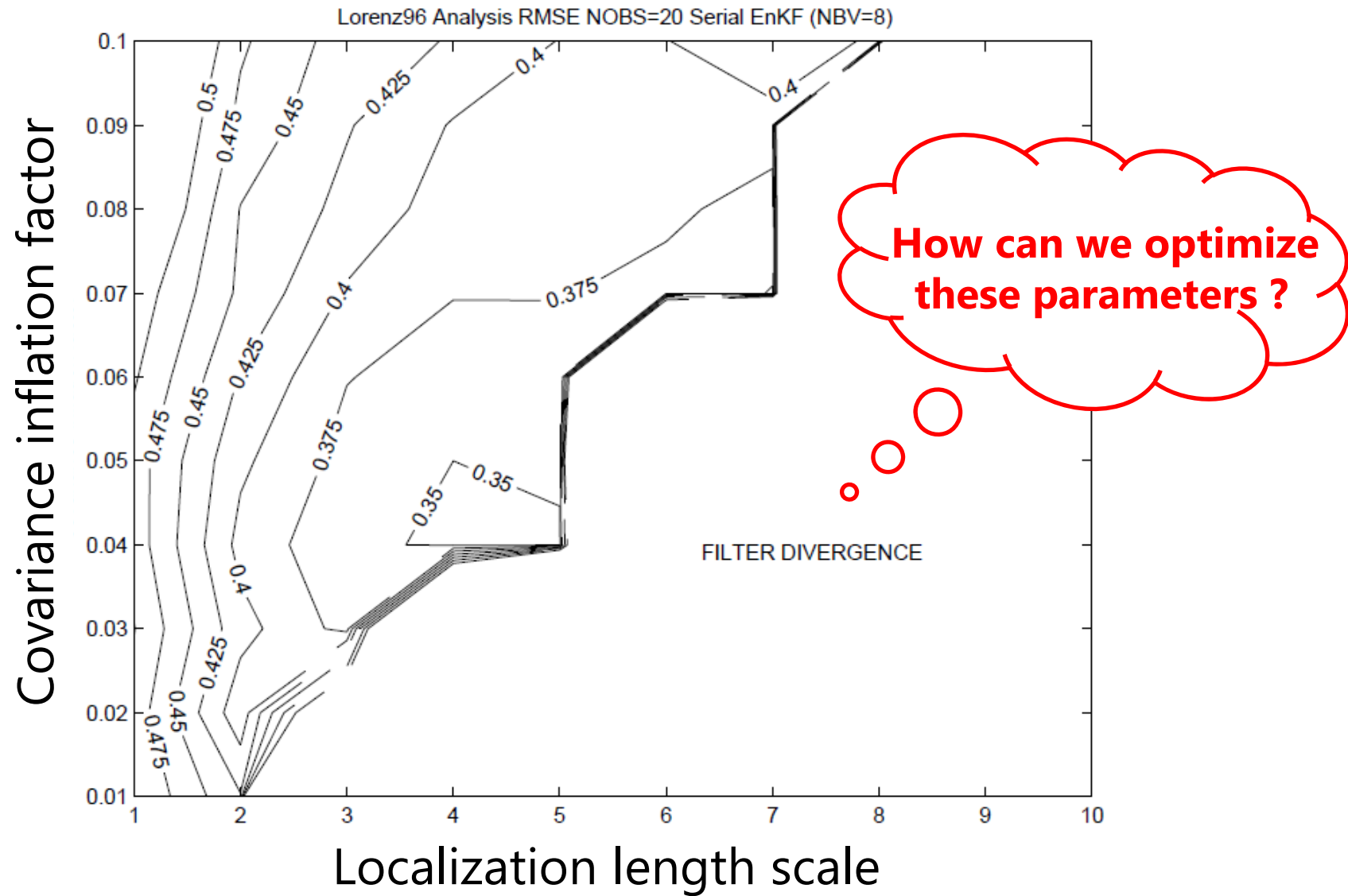
H: 観測演算子

R: 観測誤差共分散

4. 状態の解析

5. 誤差の解析

Tunable parameters of KFs/EnKFs



Miyoshi (2005), Fig. 2.7

Innovation statistics

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

b: background

$$\mathbf{d}^{o-a} = \mathbf{y}^o - \mathbf{H}\mathbf{x}^a$$

a: analysis

$$\mathbf{d}^{a-b} = \mathbf{H}\mathbf{x}^a - \mathbf{H}\mathbf{x}^b$$

o: observation

Desroziers' statistics (Desroziers et al. 2005)

$$\langle \mathbf{d}^{o-b} (\mathbf{d}^{o-b})^T \rangle = \mathbf{H}\mathbf{P}^b\mathbf{H}^T + \mathbf{R}$$

$$\langle \mathbf{d}^{a-b} (\mathbf{d}^{o-b})^T \rangle = \mathbf{H}\mathbf{P}^b\mathbf{H}^T$$

$$\langle \mathbf{d}^{o-a} (\mathbf{d}^{o-b})^T \rangle = \mathbf{R}$$

$$\langle \mathbf{d}^{a-b} (\mathbf{d}^{o-a})^T \rangle = \mathbf{H}\mathbf{P}^a\mathbf{H}^T$$

Derivation

Definition

$\langle \bullet \rangle$: 統計的期待値

$$\boldsymbol{\varepsilon}^o = \mathbf{y}^o - \mathbf{y}^t, \quad \boldsymbol{\varepsilon}^b = \mathbf{x}^b - \mathbf{x}^t, \quad \boldsymbol{\varepsilon}^a = \mathbf{x}^a - \mathbf{x}^t$$

$$\mathbf{R} = \langle \boldsymbol{\varepsilon}^o (\boldsymbol{\varepsilon}^o)^T \rangle, \quad \mathbf{P}^b = \langle \boldsymbol{\varepsilon}^b (\boldsymbol{\varepsilon}^b)^T \rangle, \quad \mathbf{P}^a = \langle \boldsymbol{\varepsilon}^a (\boldsymbol{\varepsilon}^a)^T \rangle$$

Derivation

$$\mathbf{d}^{o-b} = \mathbf{y}^o - \mathbf{H}\mathbf{x}^b = \mathbf{y}^o - \underline{\mathbf{y}^t + \mathbf{H}\mathbf{x}^t} - \mathbf{H}\mathbf{x}^b = \boldsymbol{\varepsilon}^o - \mathbf{H}\boldsymbol{\varepsilon}^b$$

$$\begin{aligned} \langle \mathbf{d}^{o-b} (\mathbf{d}^{o-b})^T \rangle &= \langle \boldsymbol{\varepsilon}^o (\boldsymbol{\varepsilon}^o)^T \rangle + \mathbf{H} \langle \boldsymbol{\varepsilon}^b (\boldsymbol{\varepsilon}^b)^T \rangle \mathbf{H}^T \\ &\quad - \cancel{\langle \boldsymbol{\varepsilon}^o (\mathbf{H}\boldsymbol{\varepsilon}^b)^T \rangle} - \cancel{\mathbf{H} \langle \boldsymbol{\varepsilon}^b (\boldsymbol{\varepsilon}^o)^T \rangle} \\ &= \mathbf{R} + \mathbf{H}\mathbf{P}^b\mathbf{H}^T \end{aligned}$$

Adaptive multiplicative inflation

Miyoshi (2011)

$$\langle \mathbf{d}^{o-b} (\mathbf{d}^{o-b})^T \rangle = \mathbf{H} \mathbf{P}^b \mathbf{H}^T + \mathbf{R}$$

$$\text{tr} \left[\langle \mathbf{d}^{o-b} (\mathbf{d}^{o-b})^T \rangle - \mathbf{R} \right] = \text{tr} \left[\mathbf{H} \mathbf{P}^b \mathbf{H}^T \right]$$

$$\Delta = \frac{\text{tr} \left[\langle \mathbf{d}^{o-b} (\mathbf{d}^{o-b})^T \rangle - \mathbf{R} \right]}{\text{tr} \left[\mathbf{H} \mathbf{P}_{est}^b \mathbf{H}^T \right]}$$

Estimated

Δ : Inflation Factor

この式で, $\mathbf{R}$ が不完全(真でない)場合はどうなるでしょうか？

Kalman filter

$$\mathbf{x}_t^f = \mathbf{M}\mathbf{x}_{t-1}^a - \boldsymbol{\eta}$$

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

Multiplicative inflation

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

$$\mathbf{x}_t^a = \mathbf{x}_t^f + \mathbf{K}_t \left[\mathbf{y}_t^o - \mathbf{H} \mathbf{x}_t^f \right]$$

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

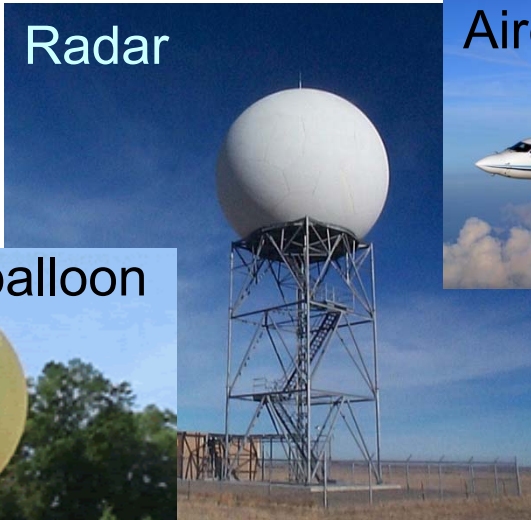
Desroziers' statistics

$$\left\langle \mathbf{d}^{o-b} (\mathbf{d}^{o-b})^T \right\rangle = \mathbf{H} \mathbf{B} \mathbf{H}^T + \mathbf{R}$$

$$\left\langle \mathbf{d}^{a-b} (\mathbf{d}^{o-b})^T \right\rangle = \mathbf{H} \mathbf{B} \mathbf{H}^T$$

Implementation: global weather forecasts

Radar



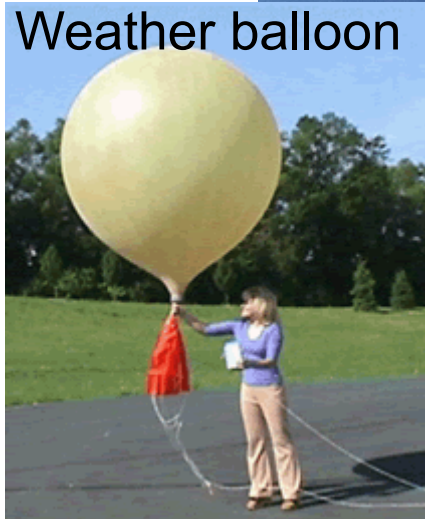
Aircraft



Satellite



Weather balloon



Ship



Buoy

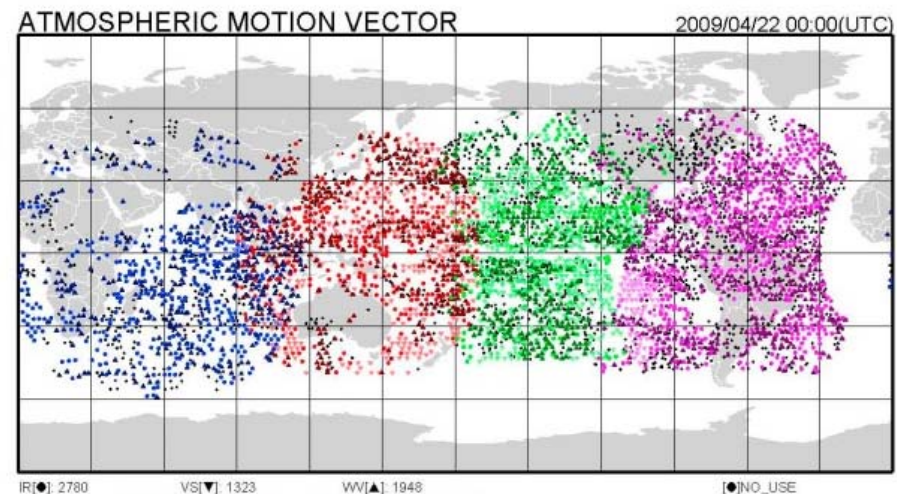
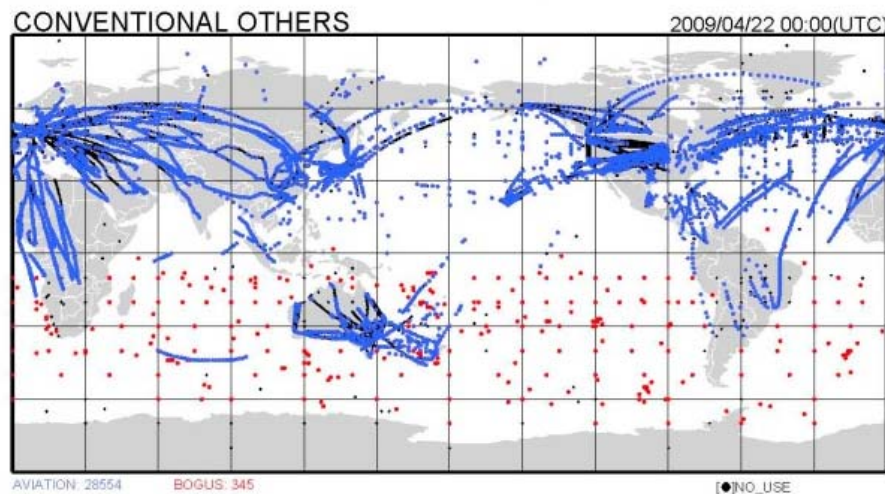
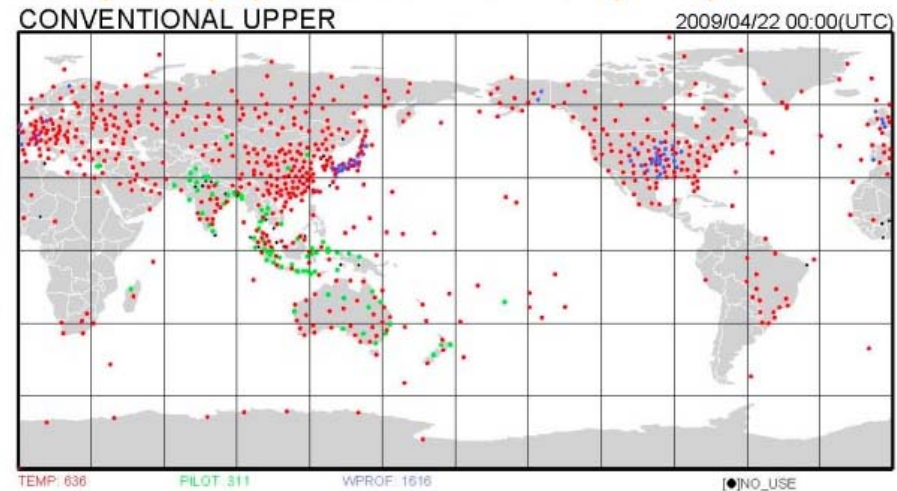
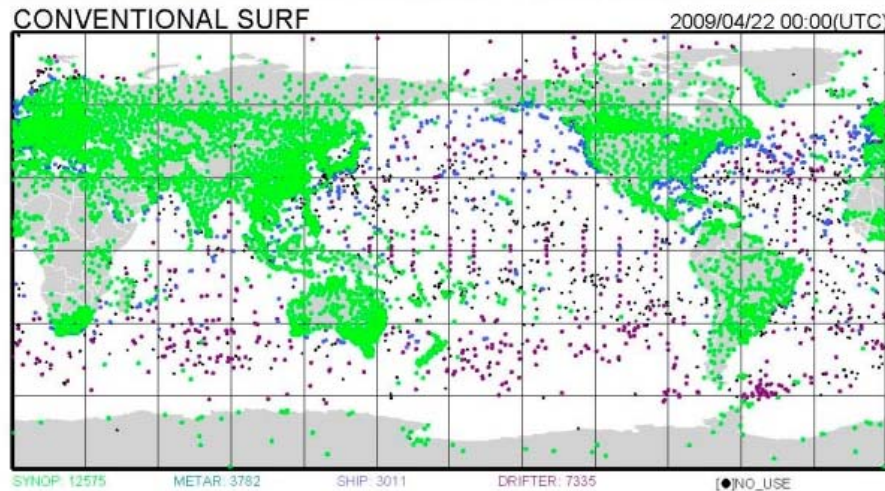


Surface station



Observation data (6-h period)

JMA GLOBAL ANALYSIS – DATA COVERAGE MAP (Da00ps): 2009/04/22 00:00(UTC)

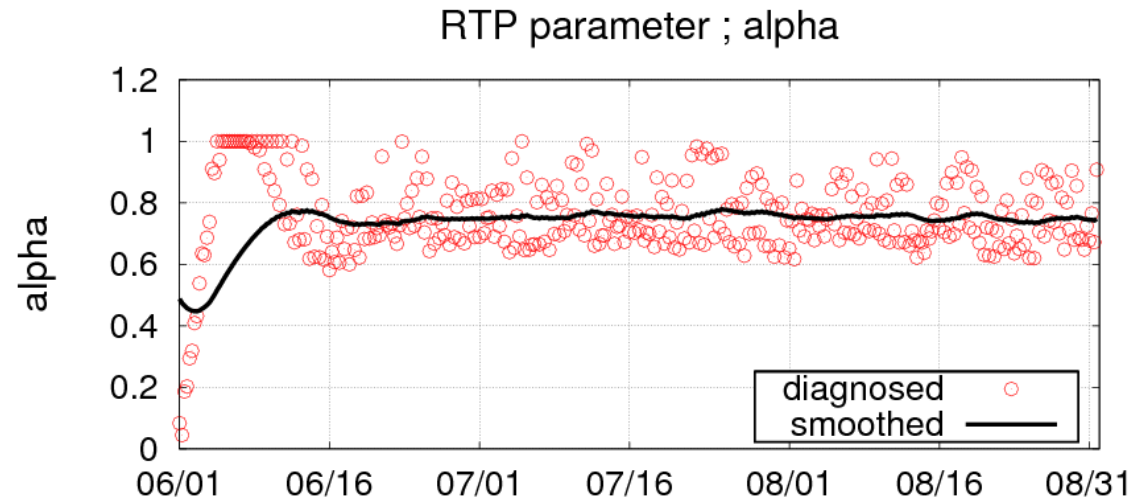


World's effort! (no border in the atmosphere)

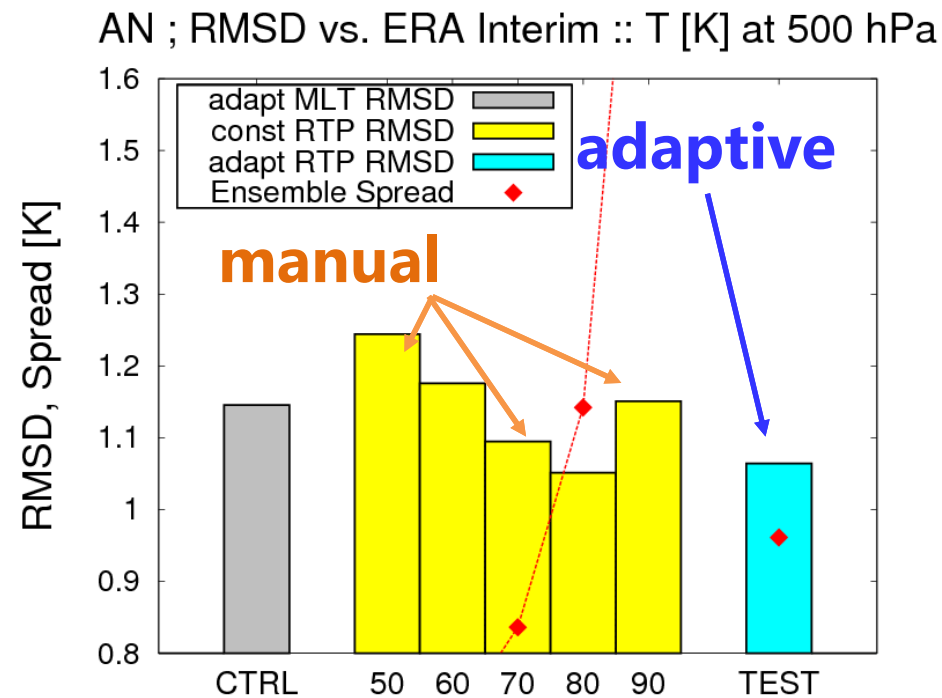
(Courtesy of JMA)

A case for Relaxation w/ NICAM-LETKF

**Adaptively
estimated α**



Change in Error



Summary

- Review: covariance inflation methods
- Understand innovation statistics
 - Desroziers et al. (2005)
- Understand adaptive inflation methods
 - Multiplicative: Miyoshi (2011), Li et al. (2009)
 - Relaxation to prior spread: Ying and Zhang (2015)
 - Relaxation to prior perturbation: Kotsuki et al. (prep)