

第4回EFD/CFD融合ワークショップ
2011年1月25日、秋葉原コンベンションホール

気象学における 4次元変分法データ同化

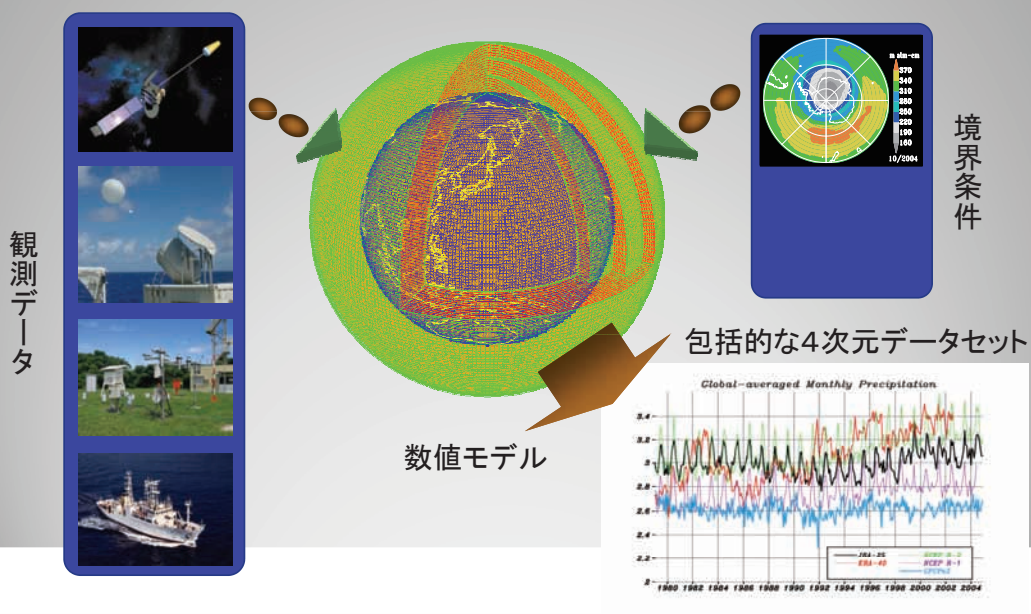
気象研究所予報研究部
露木 義

講演の内容

1. データ同化と数値天気予報
2. 4次元変分法によるデータ同化
3. 変分法とカルマンフィルタ
4. データ同化の課題

1. データ同化と数値天気予報

データ同化 観測と数値シミュレーションの融合手段



データ同化の目的

- **対象**
 - 時間発展する大規模かつ複雑なシステム
(大気、海洋、・・・)
- **目的**
 - 包括的な4次元データセットの作成
 - ➡ 研究の基盤的データ
 - 数値モデルの初期値の推定
 - ➡ 数値天気予報、数値シミュレーション
 - 数値モデルのパラメータの推定
 - ➡ 経験的定数、モデル誤差、発生・吸収源
 - 観測システムの評価・設計 など

データ同化の歴史（1）

データ同化は数値天気予報(NWP)から始まった

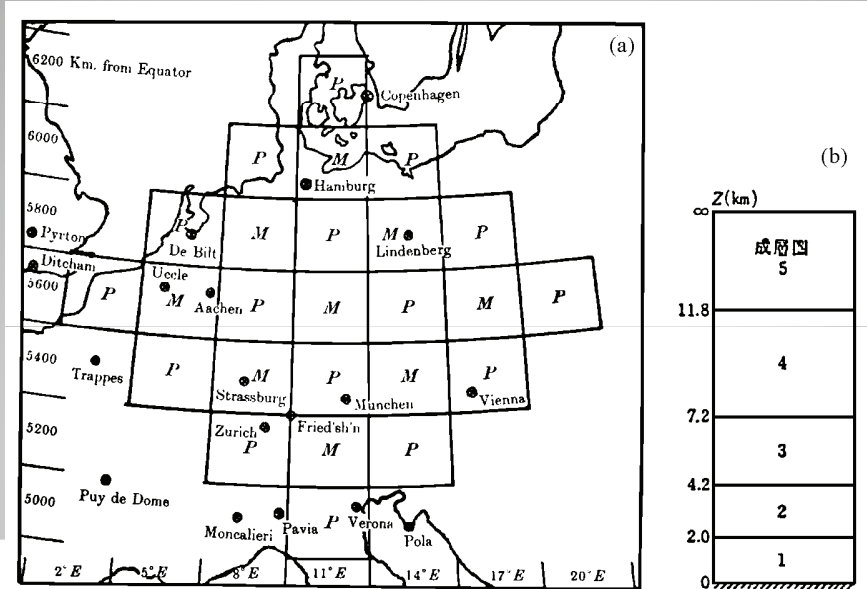
- **1922** Richardsonが手計算で数値天気予報を試みたが失敗（145hPa/6hの気圧変化を予測）
- **1950** Charneyらが電子計算機ENIACにより数値予報実験に成功（1日予報に1日要した）

初期値：主観（手書き）解析から客観（計算機）解析へ

- **1954** スウェーデン気象水理局が数値予報業務開始
- **1959** 気象庁が数値予報業務開始（世界で3番目）

客観解析法：関数あてはめ法、逐次修正法、最適内挿法
(後二者では、モデル予報値を第一推定値に採用)

Richardsonの実験の格子と観測点 (1910年5月20日07UTC)



Richardson (1922)

データ同化の歴史 (2)

非定時観測データへの対応

- **1969** Charneyらがデータ同化を提唱
- **1979** 全球大気研究計画第1回全球実験
⇒ 全球大気観測システムの確立

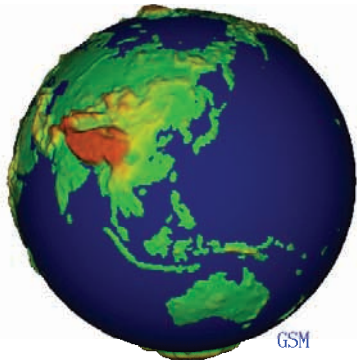
データ同化の発展

- **1980年代～** 変分法やカルマンフィルタの研究
- **1990年代～** 全球大気長期再解析の実施

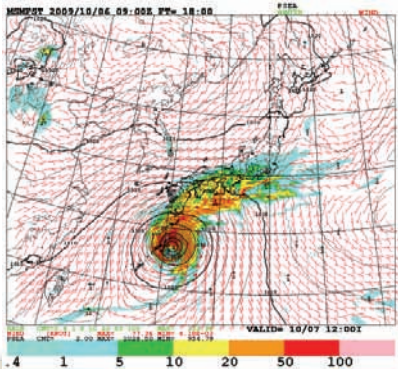
他の分野では、たとえば

- **1980年代～** 海洋データ同化の研究
- **1990年代後半～** 微量物質輸送データ同化の研究

気象庁の主な数値予報モデル



GSM



全球モデル
 水平20 km 鉛直60層
 9日予報
 天気予報、週間予報など

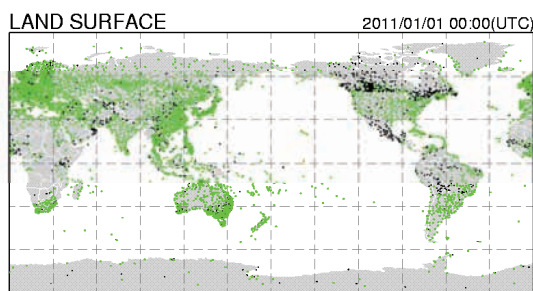


側面境界条件

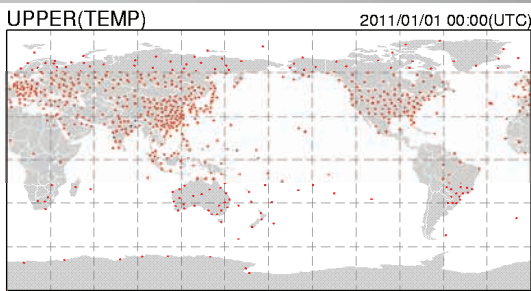
メソモデル
 水平5 km 鉛直50層
 33時間予報
 防災気象情報、航空予報など

データ同化はいずれも4次元変分法(4DVar)

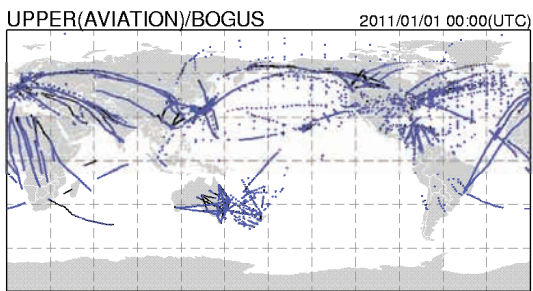
気象庁全球解析の観測データ (1)



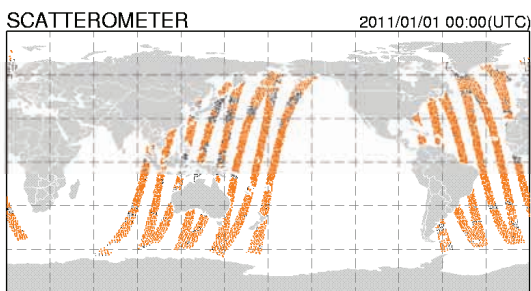
SYNOP(●) 3923 METAR
 NOUSE(●) 11247 NOUSE(●) 3749
 ALL: 15170 ALL: 3749



TEMP(●) 628
 NOUSE(●) 6
 ALL: 634

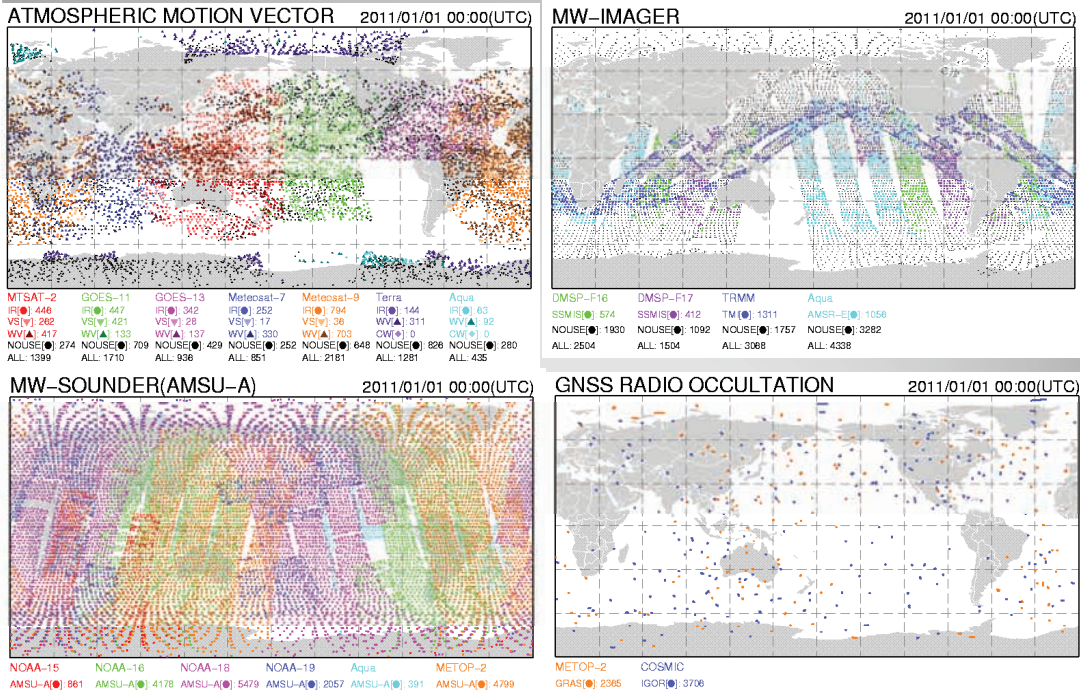


TYBOGUS(●) 0 YHTC AVIATION(●) 3939
 NOUSE(●) 0 NOUSE(▼) 15427
 ALL: 0 ALL: 0 ALL: 19386



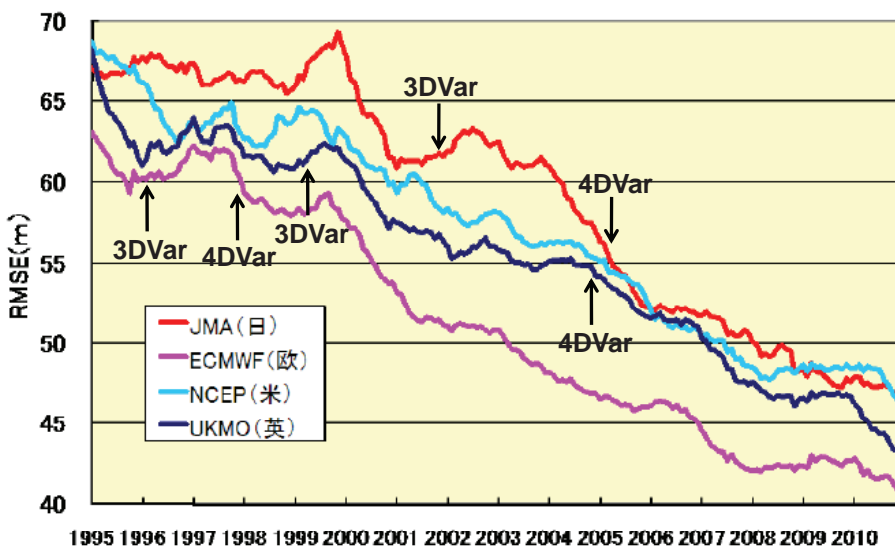
METOP-2
 ASCAT(●) 4967
 NOUSE(●) 932
 ALL: 5979

気象庁全球解析の観測データ (2)



NWPセンターの予報誤差の比較

図6-3 2010年11月までの北半球500hPa高度予報誤差(120時間予報)



2. 4次元変分法によるデータ同化

4DVar (強拘束)

評価関数 (予報誤差と観測誤差がガウス分布する場合)

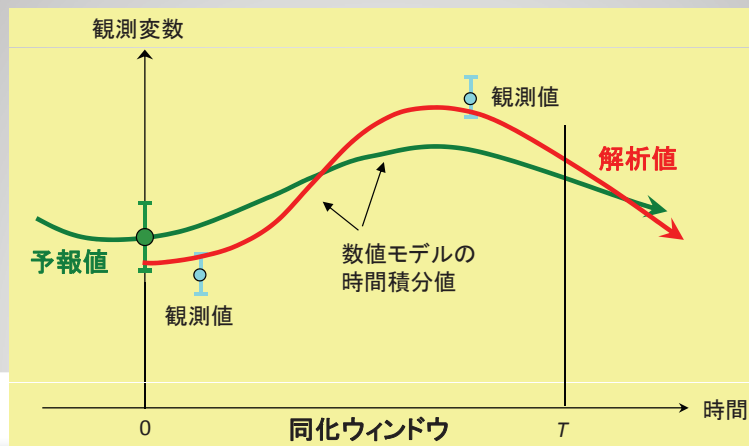
$$J(\mathbf{x}_0) = \frac{1}{2}(\mathbf{x}_0 - \mathbf{x}_0^f)^T \mathbf{B}^{-1}(\mathbf{x}_0 - \mathbf{x}_0^f) + \frac{1}{2}[H(M(\mathbf{x}_0)) - \mathbf{y}^o]^T \mathbf{R}^{-1}[H(M(\mathbf{x}_0)) - \mathbf{y}^o]$$

M : モデルの時間推進演算子

H : 観測演算子

B : 背景(予報)誤差共分散行列

R : 観測誤差共分散行列



Adjoint 法

数値モデル
(\mathbf{c} : パラメータ)

$$\frac{d\mathbf{x}}{dt} = \mathbf{F}(t, \mathbf{x}(t); \mathbf{c})$$

評価関数

$$J(\mathbf{x}(0), \mathbf{c}) = \int_0^T f(t, \mathbf{x}(t); \mathbf{c}) dt$$

・評価関数の勾配ベクトルの計算

Adjoint モデル

$$-\frac{d\mathbf{p}}{dt} = \left(\frac{\partial \mathbf{F}}{\partial \mathbf{x}} \right)^T \mathbf{p}(t) + \frac{\partial f}{\partial \mathbf{x}} \quad (\mathbf{p}(T) = \mathbf{0})$$

勾配ベクトル

$$\nabla_{\mathbf{x}(0)} J = \mathbf{p}(0)$$

$$\nabla_{\mathbf{c}} J = \int_0^T \left(\frac{\partial \mathbf{F}}{\partial \mathbf{c}} \right)^T \mathbf{p}(t) dt + \int_0^T \frac{\partial f}{\partial \mathbf{c}} dt$$

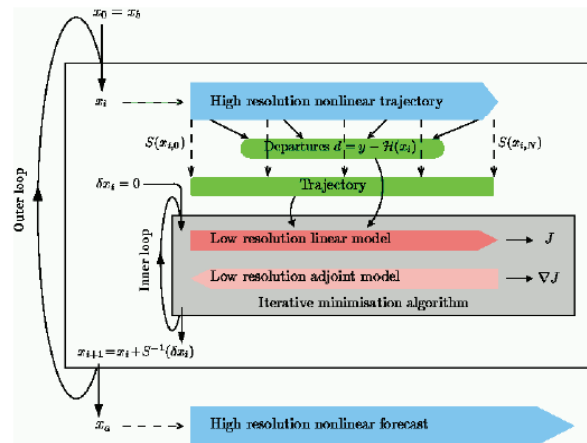
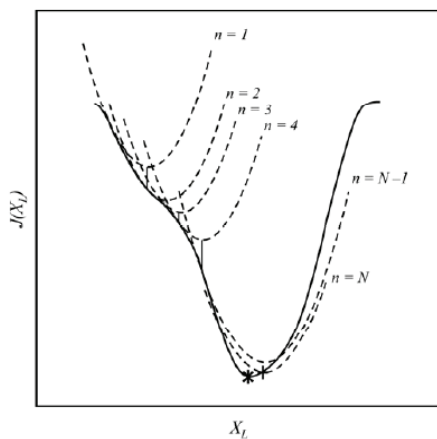
・反復法による評価関数の最小値探索

Quasi-Newton法、またはConjugate gradient法

Incremental 4DVar

計算量を大幅に減らすために・・・

- ・評価関数を局所的に2次関数で近似して反復的に解く
- ・局所的探索では、低解像度かつ単純化した線形数値モデルを用いる



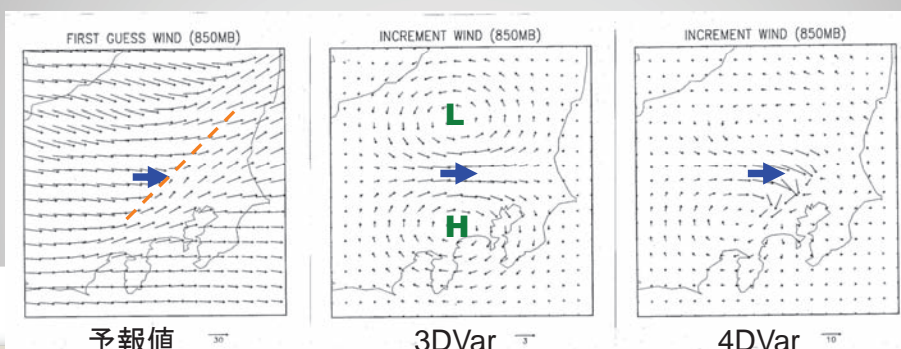
ただし、非線形性が強いとうまくいかない

Andersson et al.(2005)

従来の同化法に対する利点

| | 従来の同化法 | 4DVar |
|-----------------|----------------|---------------|
| 予報値の修正法 | 統計的知見を利用 | 数値モデルを利用 |
| 状態変数でない観測データの扱い | 状態変数に変換して同化 | そのまま同化 |
| 非定時観測データの扱い | 定時観測データとみなして同化 | 観測時刻のデータとして同化 |

風の1点観測データによる風の修正量の比較



弱拘束4DVar

モデル誤差を考慮する

数値モデルの方程式

$$dx = F(x, t)dt + \varepsilon_t$$

1. 制御変数: 初期値とモデル誤差

$$J(x_0, \{\varepsilon_t\}) = \frac{1}{2} (x_0 - x_0^f)^T B^{-1} (x_0 - x_0^f) + \frac{1}{2} [H(M(x_0, \{\varepsilon_t\})) - y^o]^T R^{-1} [H(M(x_0, \{\varepsilon_t\})) - y^o] + \frac{1}{2} \int_0^T \varepsilon_t^T Q^{-1} \varepsilon_t dt$$

2. 制御変数: 状態変数

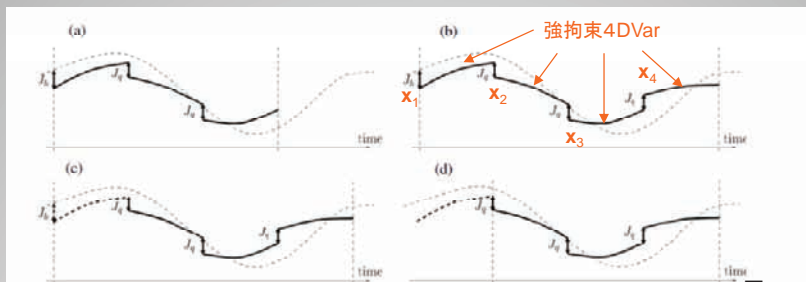
$$J(\{x_t\}) = \frac{1}{2} (x_0 - x_0^f)^T B^{-1} (x_0 - x_0^f) + \frac{1}{2} [H(M(\{x_t\})) - y^o]^T R^{-1} [H(M(\{x_t\})) - y^o] + \frac{1}{2} \int_0^T (x_t - \tilde{x}_t)^T Q^{-1} (x_t - \tilde{x}_t) dt$$

(\tilde{x}_t : モデル誤差を無視して計算した状態変数)

数値天気予報では、2009年にECMWFで初めて導入された。

Long-window 弱拘束4DVar

弱拘束にすることによって、長い同化ウィンドウを採用できる



Tremolet (2006)

評価関数

$$J(\mathbf{x}_1, \dots, \mathbf{x}_n) = J_1(\mathbf{x}_1) + \dots + J_n(\mathbf{x}_n) + J_q(\mathbf{x}_1, \mathbf{x}_2) + \dots + J_q(\mathbf{x}_{n-1}, \mathbf{x}_n)$$

$$J_q(\mathbf{x}_k, \mathbf{x}_{k+1}) := \frac{1}{2} (M(\mathbf{x}_k) - \mathbf{x}_{k+1})^T \mathbf{Q}^{-1} (M(\mathbf{x}_k) - \mathbf{x}_{k+1}), \quad (\mathbf{Q}: \text{モデル誤差共分散})$$

- $J_k(\mathbf{x}_k), \nabla J_k(\mathbf{x}_k)$ ($k=1, \dots, n$) は並列的に計算できる
- 同化ウィンドウが十分長ければ、背景誤差共分散は必要ない

Ensemble 4DVar (En4DVar)

Adjoint モデルが不要

- アンサンブル予報 (m メンバー) から背景誤差共分散を近似計算

$$\mathbf{B} = \delta \mathbf{X}^f (\delta \mathbf{X}^f)^T, \quad \delta \mathbf{X}^f := \frac{1}{\sqrt{m-1}} (\mathbf{x}_1^f - \mathbf{x}^f, \dots, \mathbf{x}_m^f - \mathbf{x}^f)$$

- 状態変数 \mathbf{x} をアンサンブル空間で近似 (\mathbf{w} : m 次元ベクトル)

$$\mathbf{x} = \mathbf{x}^f + \delta \mathbf{X}^f \mathbf{w} \quad (\langle \mathbf{w} \rangle = 0, \langle \mathbf{w} \mathbf{w}^T \rangle = \mathbf{I})$$

- 評価関数とその勾配ベクトル (制御変数: \mathbf{w})

$$J(\mathbf{w}) = \frac{1}{2} \mathbf{w}^T \mathbf{w} + \frac{1}{2} [H(M(\mathbf{x}^f + \delta \mathbf{X}^f \mathbf{w})) - \mathbf{y}^o]^T \mathbf{R}^{-1} [H(M(\mathbf{x}^f + \delta \mathbf{X}^f \mathbf{w})) - \mathbf{y}^o]$$

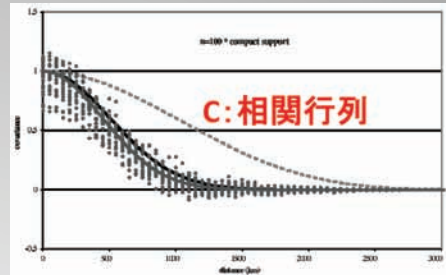
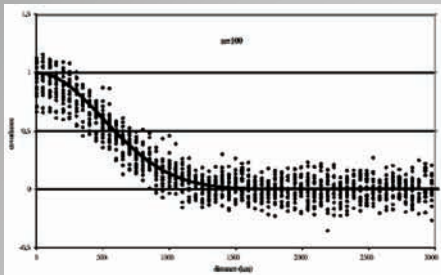
$$\nabla J(\mathbf{w}) = \mathbf{w} + \frac{1}{2} (\mathbf{H} \mathbf{M} \delta \mathbf{X}^f)^T \mathbf{R}^{-1} [H(M(\mathbf{x}^f + \delta \mathbf{X}^f \mathbf{w})) - \mathbf{y}^o]$$

- 勾配ベクトルをアンサンブル予報を使って近似計算

$$\mathbf{H} \mathbf{M} \delta \mathbf{X}^f = \frac{1}{\sqrt{m-1}} [H(M(\mathbf{x}_1^f)) - H(M(\mathbf{x}^f)), \dots, H(M(\mathbf{x}_m^f)) - H(M(\mathbf{x}^f))]$$

共分散の局所化

誤差共分散のサンプリングエラーを抑える (EnKF)



En4DVarでは・・・

Lorenz (2003)

$$\mathbf{B} = \mathbf{C} \circ (\delta \mathbf{X}^f (\delta \mathbf{X}^f)^T) \quad (\circ : \text{Schur 積})$$

$$\mathbf{x} = \mathbf{x}^f + (\text{diag}[\delta \mathbf{X}_1^f] \mathbf{C}', \dots, \text{diag}[\delta \mathbf{X}_m^f] \mathbf{C}') \mathbf{w}$$

$$\mathbf{C} = \mathbf{C}' \mathbf{C}'^T, \quad \text{diag}[\delta \mathbf{X}_k^f] := \frac{1}{\sqrt{m-1}} \begin{pmatrix} (\mathbf{x}_k^f - \mathbf{x}^f)_1 & & 0 \\ & \ddots & \\ 0 & & (\mathbf{x}_k^f - \mathbf{x}^f)_n \end{pmatrix}$$

$\mathbf{C}' : n \times r$ 行列 $\mathbf{w} : m \cdot r$ 次元

3. 変分法とカルマンフィルタ

データ同化・予測問題

- 状態変数 $\mathbf{x}(t)$ の時間発展 (β_t : Brown運動)

$$d\mathbf{x} = \mathbf{F}(\mathbf{x}, t)dt + \mathbf{G}(\mathbf{x}, t)d\beta_t, \quad \langle d\beta_t d\beta_{t'}^T \rangle = \mathbf{Q}(t)\delta(t-t')dt dt'$$

- 確率密度関数 $p(\mathbf{x}, t)$ の時間発展: **Fokker-Planck 方程式**

$$\frac{\partial p}{\partial t} + \sum_{i=1}^n \frac{\partial}{\partial x_i} (p F_i) = \frac{1}{2} \sum_{i,j=1}^n \frac{\partial^2}{\partial x_i \partial x_j} [p (\mathbf{G} \mathbf{Q} \mathbf{G}^T)_{ij}]$$

- 観測データ \mathbf{y} による $p(\mathbf{x}, t)$ の収縮: **Bayesの定理**

$$p(\mathbf{x}(t_k) | \mathbf{y}_0, \dots, \mathbf{y}_k) = \frac{p(\mathbf{y}_k | \mathbf{x}(t_k)) p(\mathbf{x}(t_k) | \mathbf{y}_0, \dots, \mathbf{y}_{k-1})}{\int p(\mathbf{y}_k | \mathbf{x}(t_k)) p(\mathbf{x}(t_k) | \mathbf{y}_0, \dots, \mathbf{y}_{k-1}) d^n \mathbf{x}(t_k)}$$

この問題を逐次モンテカルロ法で解く \Rightarrow **粒子フィルタ**
(大自由度系への適用可能性が大きな課題)

現実的なデータ同化法

確率密度関数 $p(\mathbf{x}, t)$ の平均値またはモードの直接推定
及びそれらの周りの共分散の推定

- 平均値** $\mathbf{x}^a = \langle \mathbf{x} \rangle = \int_{-\infty}^{\infty} \mathbf{x} p(\mathbf{x} | \mathbf{x}^f, \mathbf{y}^o) d\mathbf{x}$

線形最小分散推定 \Rightarrow **カルマンフィルタ** (Kalman, 1960)

$$\mathbf{x}^a = \mathbf{x}^f + \mathbf{K}(\mathbf{y}^o - H(\mathbf{x}^f)), \quad \mathbf{K} \equiv \mathbf{P}^f \mathbf{H}^T (\mathbf{R} + \mathbf{H} \mathbf{P}^f \mathbf{H}^T)^{-1}$$

Evensen (1994) : **アンサンブルカルマンフィルタ(EnKF)** の提案

- モード** $\mathbf{x}^a = \mathbf{x}_m = \arg \max_{\mathbf{x}} p(\mathbf{x} | \mathbf{x}^f, \mathbf{y}^o)$

MAP推定または最尤推定 \Rightarrow **変分法解析** (Sasaki, 1958)

$$\text{Minimize: } J(\mathbf{x}) = -\log p(\mathbf{x}^f | \mathbf{x}) - \log p(\mathbf{y}^o | \mathbf{x}) - \log p(\mathbf{x})$$

Lewis and Derber (1985), LeDimet and Talagrand (1986) など:
Adjoint法による**4次元変分法(4DVar)**の提案

状態変数の時間発展

数値モデルの方程式(ランダム項を無視)

$$\frac{d\mathbf{x}}{dt} = \mathbf{F}(\mathbf{x}, t)$$

•初期値が平均値の場合

$$\frac{d\langle \mathbf{x} \rangle}{dt} = \langle \mathbf{F}(\mathbf{x}, t) \rangle \neq \mathbf{F}(\langle \mathbf{x} \rangle, t)$$

モデルが非線形だと予測値は平均値でない → アンサンブル予報

•初期値がモードの場合

$$\frac{d\mathbf{x}_m}{dt} = \mathbf{F}(\mathbf{x}_m, t) + [p(p'')^{-1} \nabla(\nabla \cdot \mathbf{F})]_{\mathbf{x}=\mathbf{x}_m}, \quad p(\mathbf{x}, t): \text{確率密度関数}$$

モデルが非線形でも $\nabla(\nabla \cdot \mathbf{F}) = 0$ を満たせば、予測値もモード

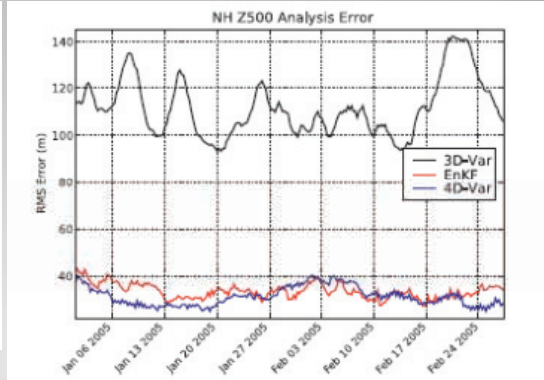
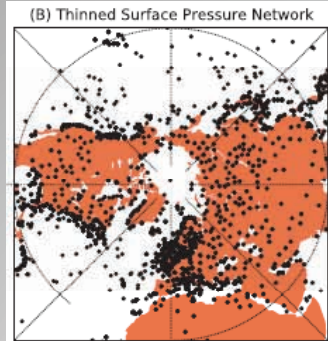
- Hamilton力学系 → 天体の運動の正確な予測?
- Lorenzモデル、順圧渦度方程式のスペクトルモデル、...

4DVarとEnKFの比較 (1)

| | 4DVar | EnKF |
|-----------|---|---|
| 統計的推定法 | MAP推定(最尤推定) | 線形最小分散推定 |
| 定式化 | 非ガウス分布や非線形に対応 | ガウス分布と線形を仮定 |
| 解析誤差共分散 | 別途、要計算 | 自然に得られる |
| システム構築の時間 | <ul style="list-style-type: none"> ・数値モデルと観測演算子のアジョイント・コードの作成 ・長期平均予報誤差共分散行列の設計 | <ul style="list-style-type: none"> ・サンプリングエラーやフィルタの発散などを抑えるためのチューニング |
| 時間発展 | 決定論的予報 | アンサンブル予報 |

4DVarとEnKFの比較 (2)

1930年代
相当の地上
観測点



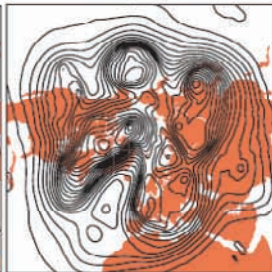
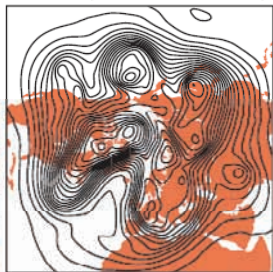
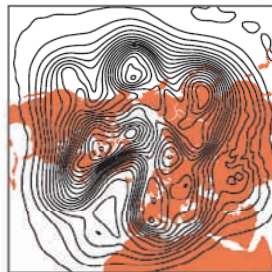
地上気圧データのみ同化

EnsDA (RMS Error = 31 m)

4D-Var (RMS Error = 31 m)

NCEP Operational

500hPa
高度



Whitaker et al. (2009)

高度なデータ同化法の現状

-変分法

- > 4DVar (strong constraint, weak constraint)
- > En4DVar

-カルマンフィルタ

- > Perturbed observations EnKF
- > Square root EnKF (EnSRF, EAKF, LETKF)

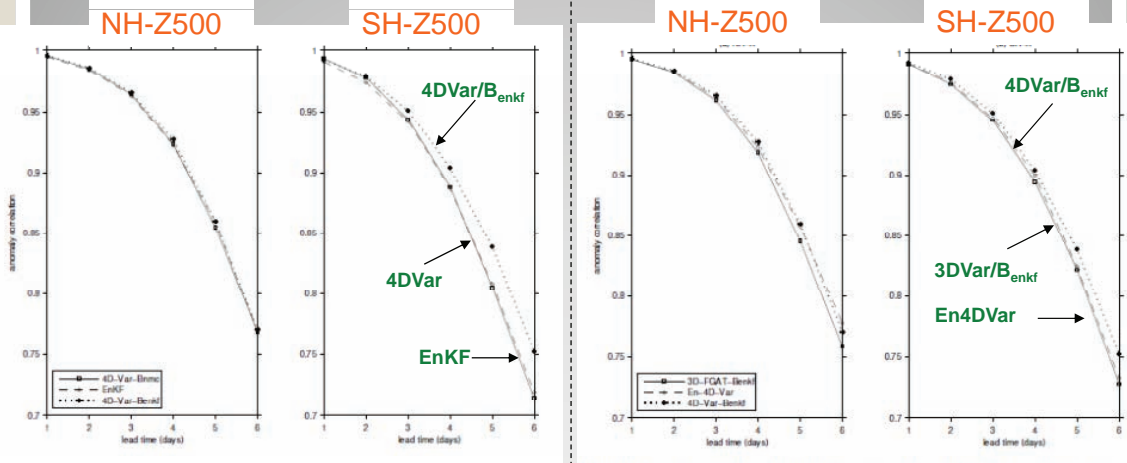
-ハイブリッド

- > 3DVar + EnKF
- > 4DVar + EnKF

-粒子フィルタ

データ同化法の比較（予報精度）

カナダ気象局の全球モデル(400×200×58L)：2007年2月



- EnKFと4DVarは同程度の精度
- 4DVarや3DVarにおいてEnKFで計算したBを使う(ハイブリッド)と改善
- En4DVarは、EnKFや4DVarとほぼ同程度の精度

Buehner et al. (2010)

4. データ同化の課題

1. 非線形と non-Gaussian

- モデルの非線形性の強さと観測データの時空間密度との兼ね合いに依存する。
- 集中豪雨の予測: モデルの雲物理過程は非線形性が強く、観測データは不十分。



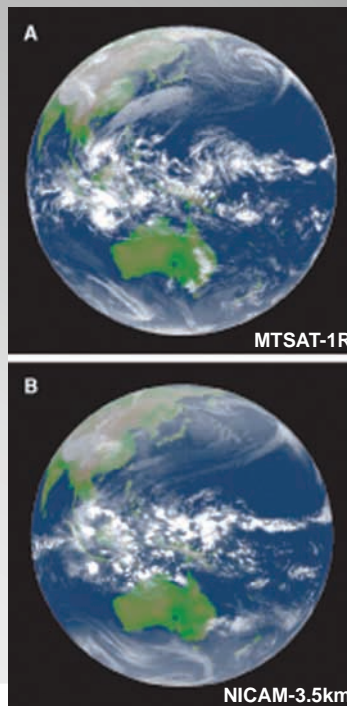
Long window 弱拘束4DVar、Incremental EnKS、粒子フィルタとのハイブリッド、...

2. マルチスケール

- 大気海洋結合モデル: 大気と海洋の時間スケールの違い
- 集中豪雨の予測: 積乱雲(1~10km)と環境場(10²~10³km)の時空間スケールの違い。



- 階層的アプローチ、...



H Miura et al. Science 2007;318:1763-1765

次世代スパコン戦略プログラム

分野3: 防災・減災に資する地球変動予測

(平成23~27年度) 参加機関: JAMSTEC、東京大学、気象研究所など

背景

地球温暖化時の台風の動向が今なお不明確、より高精度の集中豪雨予測のニーズ
地震発生時における避難・救援行動に必要な高解像度の被害予測、津波による浸水域等の高精度な情報のニーズ

戦略目標

地球温暖化時の台風の動向の全球的予測と**集中豪雨の予測実証**
および次世代型**地震**ハザードマップの構築と**津波**警報の高精度化

研究開発課題

(1) 防災・減災に資する気象・気候・環境予測研究

(2) 地震・津波の予測精度の高度化に関する研究

