

最尤法アンサンブルフィルタを用いた非線形観測の同化

Assimilation of Nonlinear Observations Using the Maximum Likelihood Ensemble Filter

中下 早織⁽¹⁾・榎本 剛

Saori NAKASHITA⁽¹⁾ and Takeshi ENOMOTO

(1) 京都大学大学院 理学研究科

(1) Graduate School of Science, Kyoto University, Japan

Synopsis

We investigate the performance of the Maximum Likelihood Ensemble Filter (MLEF) in assimilation of nonlinear observations. MLEF is a variational-ensemble data assimilation method, and can treat differentiable or non-differentiable nonlinear observation operators. In this study, we compare MLEF with the Ensemble Transform Kalman Filter (ETKF) in assimilation experiments with a one-dimensional Burgers model. The ETKF analysis with a certain formulation of nonlinear operators diverges when the observation nonlinearity is strong and the observation error is small. This divergence is found to be associated with an extra rank of ensemble perturbation matrix. Optimization in MLEF can improve the analysis to the level comparable to or better than ETKF. In addition, the smaller observation error is, or the stronger observation nonlinearity is, MLEF with the nonlinear operators can assimilate observations more effectively than MLEF with the tangent linear operators. However, the strong nonlinearity hinders convergence. We found that re-evaluation of the Hessian preconditioning matrix can alleviate such poor convergence. These encouraging results indicate that MLEF can incorporate nonlinear effects and evaluate observations appropriately.

キーワード : 最尤推定, アンサンブルデータ同化, 数値最適化, 非線形観測演算子

Keywords : Maximum Likelihood Estimation, Ensemble data assimilation, Numerical optimization, Nonlinear observation operator

1. はじめに

気象予報は気象をシミュレーションする予報モデルと観測と予報を組み合わせるデータ同化システムによって作られる。予報精度向上のためには、予報モデルの高解像度化・精緻化とともに、より多くの観測の情報を同化によって取り入れることが重要である。気象要素の観測の中で近年利用が増えているのは、衛星による観測である。衛星観測は、陸上に観測点が集中する地上観測や観測時刻が限定的である高層観測と比較して、空間・時間ともに密な観測が得

られることが利点である。一方で、衛星によって観測される要素は予報モデルで扱われる要素とは異なるため変換が必要となり同化が複雑になること、観測間の相関や観測のバイアスを考慮する必要があることが克服すべき点である。特に観測要素と予報要素間の変換を表す観測演算子は非線形であり、この非線形性が大きな課題となる。例えば衛星によって観測される放射輝度 α を輝度温度 T に変換するシュテファン・ボルツマンの法則 $\alpha = \sigma T^4$ (σ はシュテファン・ボルツマン定数) は 4 次の非線形性を持つ。

多くのデータ同化手法は観測演算子を予報変数の近傍で接線形近似して扱うため、非線形の観測演算子は近似的にしか扱うことができない。また、水蒸気関連量を扱う場合には雲の有無によって関数を切り替える必要があり、観測演算子に不連続性が含まれ微分不可能となる。多くの現業予報機関で用いられる変分法は観測演算子の微分可能性を仮定して導出されるため、この不連続性も障害となる。

観測演算子の線形近似は、アンサンブルを利用したデータ同化手法を用いると避けることができる。アンサンブルデータ同化手法として良く知られているのはアンサンブルカルマンフィルタ (Evensen 1994) であるが、アンサンブルカルマンフィルタは線形の枠組みで導出されるカルマンフィルタを基礎としているため、強い非線形を含む観測演算子は想定されていない。また不連続性への適応も定かではない。そこで本研究では、アンサンブルと変分法を組み合わせた最尤法アンサンブルフィルタ (MLEF, Zupanski 2005) を非線形観測の同化に適用する。MLEF は変分法で扱われるコスト関数をアンサンブル空間で解くことで解析値を求める同化手法であり、線形でも非線形でも同じ定式化で同化することができることに加え、微分不可能な演算子も扱うことができる (Zupanski et al. 2008)。本研究の目的は、将来衛星観測の同化に応用することを念頭に、MLEF の非線形観測同化の性能を単純化した実験で検証することである。特に観測演算子の接線形近似の有無による解析精度の違いに注目して結果を分析する。また、非線形演算子を扱えるとされるアンサンブルカルマンフィルタの中でも MLEF と似た定式化を持つアンサンブル変換カルマンフィルタ (ETKF, Bishop et al. 2001) との比較を行う。

2. 同化手法

この章では本研究で用いる 2 つの同化手法について説明する。以下の導出に共通する数学記号を定義する。

- n 状態変数の次元
- p 観測数
- m アンサンブルサイズ
- $\mathbf{x}^f \in \mathbb{R}^n$ 予報値
- $\mathbf{x}^a \in \mathbb{R}^n$ 解析値
- $\mathbf{y} \in \mathbb{R}^p$ 観測値

- $\mathbf{P}_f \in \mathbb{R}^{n \times n}$ 予報誤差共分散
- $\mathbf{P}_a \in \mathbb{R}^{n \times n}$ 解析誤差共分散
- $\mathbf{R} \in \mathbb{R}^{p \times p}$ 観測誤差共分散
- $H: \mathbb{R}^n \mapsto \mathbb{R}^p$ 観測演算子
- $\mathbf{H} \in \mathbb{R}^{p \times n}$ 観測演算子の接線形演算子 ($= \frac{\partial H}{\partial \mathbf{x}}$)
- $M: \mathbb{R}^n \mapsto \mathbb{R}^n$ 予報モデル
- $\mathbf{M} \in \mathbb{R}^{n \times n}$ 予報モデルの接線形演算子 ($= \frac{\partial M}{\partial \mathbf{x}}$)
- $\mathbf{I}_{p,m}$ 単位行列 (下付き添え字が次数を表す)

2.1 MLEF

MLEF ではコントロールラン (アンサンブルメンバーとは独立な予報) に対して解析を行う。以下の導出は Zupanski(2005) に従う。

予報

解析誤差共分散が以下のように分解できるとする。

$$\mathbf{P}_a = \mathbf{P}_a^{1/2} \mathbf{P}_a^{T/2}$$

$$\mathbf{P}_a^{1/2} = (\mathbf{p}_1^a, \dots, \mathbf{p}_m^a), \quad \mathbf{p}_i^a \in \mathbb{R}^n$$

ただし、 $\mathbf{P}_a^{T/2} = (\mathbf{P}_a^{1/2})^T$ である。アンサンブルメンバーを平方根行列 $\mathbf{P}_a^{1/2}$ の列ベクトル \mathbf{p}_i^a とコントロールランの和として作成する。

$$\mathbf{x}_i^a = \mathbf{x}^a + \mathbf{p}_i^a \quad (1)$$

コントロールランとアンサンブルに対して予報を行い、予報値と予報誤差共分散の平方根行列を作成する。

$$\mathbf{x}^f = M(\mathbf{x}^a) \quad (2)$$

$$\mathbf{x}_i^f = M(\mathbf{x}_i^a) \quad \text{for } i = 1, \dots, m \quad (3)$$

$$\mathbf{P}_f^{1/2} = (\mathbf{p}_1^f, \dots, \mathbf{p}_m^f)$$

$$= (\mathbf{x}_1^f - \mathbf{x}^f, \dots, \mathbf{x}_m^f - \mathbf{x}^f) \quad (4)$$

式 (4) はカルマンフィルタにおける誤差共分散の時間発展

$$\mathbf{P}_f = \mathbf{M} \mathbf{P}_a \mathbf{M}^T$$

のアンサンブル近似になっている。

解析

解析値はコスト関数

$$J(\mathbf{x}) = \frac{1}{2} (\mathbf{x} - \mathbf{x}^f)^T \mathbf{P}_f^{-1} (\mathbf{x} - \mathbf{x}^f)$$

$$+ \frac{1}{2} [\mathbf{y} - H(\mathbf{x})]^T \mathbf{R}^{-1} [\mathbf{y} - H(\mathbf{x})] \quad (5)$$

を最小化する値として求める。

解析値と予報値の差（インクリメント）を \mathbf{p}_i^f の線形結合で表す。

$$\mathbf{x} - \mathbf{x}^f = w_1 \mathbf{p}_1^f + \cdots + w_m \mathbf{p}_m^f = \mathbf{P}_f^{1/2} \mathbf{w} \quad (6)$$

$$\mathbf{w} = (w_1, \dots, w_m)^T \in \mathbb{R}^m$$

式 (6) を式 (5) に代入すると、係数 \mathbf{w} に対するコスト関数

$$J(\mathbf{w}) = \frac{1}{2} \mathbf{w}^T \mathbf{w} + \frac{1}{2} [\mathbf{y} - H(\mathbf{x})]^T \mathbf{R}^{-1} [\mathbf{y} - H(\mathbf{x})] \quad (7)$$

$$\mathbf{x} = \mathbf{x}^f + \mathbf{P}_f^{1/2} \mathbf{w}$$

を得る。最小化の際に用いるコスト関数の勾配は

$$\nabla_{\mathbf{w}} J = \mathbf{w} + (\mathbf{R}^{-1/2} \mathbf{H} \mathbf{P}_f^{1/2})^T \mathbf{R}^{-1/2} [\mathbf{y} - H(\mathbf{x}^f + \mathbf{P}_f^{1/2} \mathbf{w})] \quad (8)$$

で与えられる。ここで、行列 $\mathbf{Z}(\mathbf{x})$ を以下で定義する。

$$\mathbf{Z}(\mathbf{x}) = (\mathbf{z}_1, \dots, \mathbf{z}_m) \in \mathbb{R}^{p \times m}$$

$$\mathbf{z}_i = [\mathbf{R}^{-1/2} \mathbf{H} \mathbf{P}_f^{1/2}]_i \quad (9)$$

$$\simeq \mathbf{R}^{-1/2} [H(\mathbf{x} + \mathbf{p}_i^f) - H(\mathbf{x})] \quad (10)$$

最終行は観測演算子のテイラー展開を利用した拡張であり、接線形近似 \mathbf{H} を避けた導出となっている。この行列 $\mathbf{Z}(\mathbf{x})$ を用いると、勾配 (8) およびヘシアンを以下のように導出できる。

$$\nabla_{\mathbf{w}} J = \mathbf{w} + \mathbf{Z}(\mathbf{x})^T \mathbf{R}^{-1/2} [\mathbf{y} - H(\mathbf{x})] \quad (11)$$

$$\nabla_{\mathbf{w}}^2 J = \mathbf{I}_m + \mathbf{C}(\mathbf{x}) \quad (12)$$

ここで、 $\mathbf{C} = \mathbf{Z}(\mathbf{x})^T \mathbf{Z}(\mathbf{x})$ である。

さらに、最小化の収束を早めるために、ヘシアンを利用した前処理（preconditioning）を導入する。

$$\mathbf{w} = [\mathbf{I}_m + \mathbf{C}(\mathbf{x}^f)]^{-1/2} \boldsymbol{\zeta} \quad (13)$$

前処理に出てくる行列は、以下のように固有値分解を利用して求めることが出来る。

$$\mathbf{C}(\mathbf{x}^f) = \mathbf{V} \boldsymbol{\Lambda} \mathbf{V}^T \quad (14)$$

$$[\mathbf{I}_m + \mathbf{C}(\mathbf{x}^f)]^{-1/2} = \mathbf{V} [\mathbf{I}_m + \boldsymbol{\Lambda}]^{-1/2} \mathbf{V}^T \quad (15)$$

コスト関数の変分 δJ を考えると、

$$\delta J = \nabla_{\mathbf{w}} J^T \delta \mathbf{w} = \nabla_{\boldsymbol{\zeta}} J^T \delta \boldsymbol{\zeta}$$

より、変数変換 (13) と合わせて

$$\nabla_{\boldsymbol{\zeta}} J = [\mathbf{I}_m + \mathbf{C}(\mathbf{x}^f)]^{-T/2} \nabla_{\mathbf{w}} J$$

と表せるため、 $\boldsymbol{\zeta}$ に対するコスト関数と勾配は以下のように表せる。

$$J(\boldsymbol{\zeta}) = \frac{1}{2} \boldsymbol{\zeta}^T [\mathbf{I}_m + \mathbf{C}(\mathbf{x}^f)]^{-1} \boldsymbol{\zeta} + \frac{1}{2} [\mathbf{y} - H(\mathbf{x})]^T \mathbf{R}^{-1} [\mathbf{y} - H(\mathbf{x})] \quad (16)$$

$$\nabla_{\boldsymbol{\zeta}} J = [\mathbf{I}_m + \mathbf{C}(\mathbf{x}^f)]^{-1} \boldsymbol{\zeta} - [\mathbf{I}_m + \mathbf{C}(\mathbf{x}^f)]^{-T/2} \mathbf{Z}(\mathbf{x})^T \mathbf{R}^{-1/2} [\mathbf{y} - H(\mathbf{x})] \quad (17)$$

$$\mathbf{x} = \mathbf{x}^f + \mathbf{P}_f^{1/2} [\mathbf{I}_m + \mathbf{C}(\mathbf{x}^f)]^{-1/2} \boldsymbol{\zeta}$$

上記のように導出したコスト関数を最小化する値 $\boldsymbol{\zeta}_{\text{opt}}$ を数値最適化手法を利用して求める。本研究では特に断らない限り、最適化手法として非線形共役勾配法（Fletcher and Reeves 1964）を用いる。解析値および解析誤差共分散の平方根行列は以下のように計算する。

$$\mathbf{x}^a = \mathbf{x}^f + \mathbf{P}_f^{1/2} [\mathbf{I}_m + \mathbf{C}(\mathbf{x}^f)]^{-1/2} \boldsymbol{\zeta}_{\text{opt}} \quad (18)$$

$$\mathbf{P}_a^{1/2} = \mathbf{P}_f^{1/2} [\mathbf{I}_m + \mathbf{C}(\mathbf{x}^a)]^{-1/2} \quad (19)$$

2.2 ETKF

ETKF ではアンサンブル平均に対して解析を行う。

予報

予報誤差共分散をアンサンブル予報値のサンプル共分散で近似する。

$$\mathbf{x}_i^f = M(\mathbf{x}_i^a) \quad \text{for } i = 1, \dots, m \quad (20)$$

$$\mathbf{P}_f = \mathbf{X}^f (\mathbf{X}^f)^T \quad (21)$$

$$\mathbf{X}^f = \frac{1}{\sqrt{m-1}} [\mathbf{x}_1^f - \overline{\mathbf{x}^f}, \dots, \mathbf{x}_m^f - \overline{\mathbf{x}^f}] \quad (22)$$

$\overline{(\cdot)}$ はアンサンブル平均を表す。

解析

アンサンブル平均の解析値を、予報のアンサンブル平均と観測の重み付き平均で求める。

$$\overline{\mathbf{x}^a} = \overline{\mathbf{x}^f} + \mathbf{K}(\mathbf{y} - \overline{H(\mathbf{x}^f)}) \quad (23)$$

ここで、重みを決める行列 \mathbf{K} （カルマンゲイン）は以下のように計算される。

$$\mathbf{K} = \mathbf{P}_f \mathbf{H}^T (\mathbf{H} \mathbf{P}_f \mathbf{H}^T + \mathbf{R})^{-1} \simeq \mathbf{X}^f \mathbf{Z}^T (\mathbf{Z} \mathbf{Z}^T + \mathbf{I}_p)^{-1} \mathbf{R}^{-1/2} = \mathbf{X}^f (\mathbf{I}_m + \mathbf{Z}^T \mathbf{Z})^{-1} \mathbf{Z}^T \mathbf{R}^{-1/2} \quad (24)$$

最終行への変形は Sherman-Morisson-Woodbery の公式を適用した等式

$$(\mathbf{I}_p + \mathbf{Z} \mathbf{Z}^T)^{-1} = \mathbf{I}_p - \mathbf{Z} (\mathbf{I}_m + \mathbf{Z}^T \mathbf{Z})^{-1} \mathbf{Z}^T$$

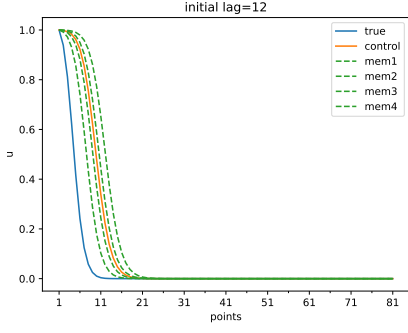


Fig. 1: Initial state for truth (blue) and experiment (control:orange, ensemble:green dashed)

を代入することで得られる．行列 \mathbf{Z} は MLEF と同様に定義され，接線形近似を避けて求めることができる．

$$\mathbf{Z} = \mathbf{R}^{-1/2} \mathbf{H} \mathbf{X}^f \quad (25)$$

$$\simeq \frac{1}{\sqrt{m-1}} \mathbf{R}^{-1/2} [H(\mathbf{x}_i^f) - \overline{H(\mathbf{x}^f)}]_{i=1, \dots, m} \quad (26)$$

$$\simeq \frac{1}{\sqrt{m-1}} \mathbf{R}^{-1/2} [H(\mathbf{x}_i^f) - H(\overline{\mathbf{x}^f})]_{i=1, \dots, m} \quad (27)$$

ただし，平均をとる操作と観測演算子を作用させる操作の順番によって上記の2通りの導出方法がある．

アンサンブルメンバーの解析値は，式 (23) で求めた平均の解析値に予報アンサンブル摂動を変換して求めた解析アンサンブル摂動を加えることで生成する．

$$\mathbf{x}_i^a = \overline{\mathbf{x}^a} + \mathbf{X}_i^a \quad (28)$$

$$\mathbf{X}^a = \mathbf{X}^f \mathbf{T}$$

$$\mathbf{T} = (\mathbf{I}_m + \mathbf{Z}^T \mathbf{Z})^{-1/2} \quad (29)$$

最終行の平方根行列は式 (14), (15) と同様に固有値分解を利用して求めることができる．

3. 実験設定

3.1 モデル

本研究では Burgers モデル (Burgers 1948)

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} = \nu \frac{\partial^2 u}{\partial x^2} \quad (30)$$

を用いる．これは移流と拡散を含む 1 次元のモデルであり，数値予報モデルの基本形である．空間方向には等間隔に格子を取り，格子間隔は 0.05，格子点数は 81 点である．時間刻み幅は 0.0125 で，拡散係数 ν は 0.05 である．値は全て無次元量である．境界

条件は領域左端で $u = 1$ ，右端で $u = 0$ であり，この境界条件に対応する解は左から右へ伝播する衝撃波を表す．

真値と実験で異なる初期値を与える双子実験を行う．モデル誤差は考慮せず，真値の作成と実験には同じモデルを用いる．初期値の分布を Fig.1 に示す．実験のコントロールランの初期値は真値に対して 40 ステップ先に進んでおり，アンサンブルメンバーの初期値はコントロールの周りに等間隔で分布させる．

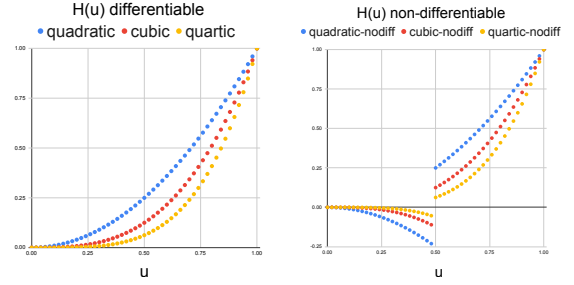


Fig. 2: Nonlinear functions for observation operators

3.2 観測

観測演算子は，非線形性の強さと不連続点の有無を変えた以下の 6 種類の関数を用いて定義する．

$$H(u) = u^k \text{ (differentiable)}$$

$$H(u) = \begin{cases} -u^k & (u < 0.5) \\ u^k & (u \geq 0.5) \end{cases} \text{ (non-differentiable)}$$

$$(k = 2, 3, 4)$$

予報値と観測演算子を作用させた値の関係を Fig.2 に示す．観測は真値に上記の観測演算子を作用させた後，平均 0 のガウス分布に従う乱数を観測誤差として加えて作成する．ガウス分布の標準偏差は $10^{-5} \sim 0.5$ の間の複数の値を用いて実験を行う．観測誤差間の相関は考慮せず，従って \mathbf{R} は観測誤差分散を対角成分に持つ対角行列とする．観測は 20 ステップごとに全ての格子点に与え，解析-予報サイクルは 20 回行う．

3.3 実験設定

以下の 5 種類の実験を行う．

MLEF-FH 同化手法として MLEF を用いて，式 (10) で \mathbf{Z} 行列を計算する．

MLEF-JH 同化手法として MLEF を用いて，式 (9) で \mathbf{Z} 行列を計算する．

ETKF-FH

av-op 同化手法として ETKF を用いて、式 (26) で \mathbf{Z} 行列を計算する。

op-av 同化手法として ETKF を用いて、式 (27) で \mathbf{Z} 行列を計算する。

ETKF-JH 同化手法として ETKF を用いて、式 (25) で \mathbf{Z} 行列を計算する。

アンサンブル数は自由度をそろえるために、MLEF で 4, ETKF で 5 とする。アンサンブル数をこれ以上大きくしても、解析精度に大きな変化はなかった。

アンサンブル同化では、サンプル数不足による解析精度の低下を緩和するために、共分散膨張 (inflation) と局所化 (localization) という 2 つの手法が良く用いられるが、今回の実験ではどちらも用いていない。

3.4 精度評価

各実験の精度評価は真値と解析値の間の RMSE

$$\text{RMSE} = \sqrt{\frac{1}{N_{\text{state}}} \sum_{i=1}^{N_{\text{state}}} [u_{\text{analysis}}(i) - u_{\text{truth}}(i)]^2} \quad (31)$$

で測る。 $N_{\text{state}} = 81$ は格子点数である。RMSE が観測誤差標準偏差よりも小さな値をとると、観測で予報を置き換えるよりも良い精度の解析値が得られたことを示す。

観測誤差の大きさを変えた影響を考慮した総合的な評価指標として、RMSE と観測誤差の比を幾何平均したスコアを導入する。

$$\text{score} = \left[\prod_{\sigma_o} \left\{ \frac{1}{N_{\text{test}}} \sum_{i=1}^{N_{\text{test}}} \left(\frac{\sigma_o}{e_{i,\sigma_o}} \right) \right\} \right]^{\frac{1}{N_{\sigma_o}}} \quad (32)$$

σ_o は観測誤差標準偏差、 e_{i,σ_o} は観測誤差 σ_o の観測を与えた i 番目の実験の RMSE (後半 15 サイクルの平均)、 N_{test} は試行回数 (本研究では 50 回)、 N_{σ_o} は観測誤差の種類数 (本研究では 15 種類) を表す。{} 内は試行の算術平均を表す。スコアが 1 と等しいとき観測誤差と解析誤差が同水準であり、大きいほど総合的に良い精度であることを示す。

4. 結果

4.1 ETKF との比較

Fig.3 に観測誤差 10^{-3} の観測を与えた場合の解析誤差の推移を示す。接線形近似を避けた実験 (FH) では、微分可能な 2 次の観測を除いて MLEF と ETKF の av-op は同水準の精度を示している。一

方で ETKF-FH op-av は微分不可能な場合の精度は av-op と同等または優れているが、微分可能な場合には精度が著しく悪く、3 次以上では発散している。微分可能な演算子 (上段) と微分不可能な演算子 (下段) を比較すると、微分不可能な場合のほうが精度が悪化しており、特に 2 次の観測で顕著である。これは Fig.2 に示す通り、2 次の微分不可能な演算子が最も不連続性が大きく、難しい問題設定となっているためであると考えられる。接線形近似を用いた実験 (JH) では、すべての演算子に対して MLEF のほうが ETKF よりも精度が良く、これは MLEF で用いるコスト関数の最適化が非線形観測同化に有効であることを示唆している。微分可能な演算子で非線形性が強い場合、JH の精度は FH に劣る。

観測誤差の大きさの影響を考慮したスコアの結果を Fig.4 に示す。Fig.3 と同様に、非線形性が強くなるほど MLEF-FH の精度がよくなっていることと、MLEF-JH がどの観測演算子においても ETKF-JH よりも優れていることがわかる。微分可能な演算子における ETKF-FH op-av は上述したように発散しているため、図には表れていない。ETKF-FH op-av は観測誤差が小さくなると計算が不安定になり、その観測誤差の閾値は非線形性が強くなるほど大きくなる (図は省略)。微分不可能な場合、すべての実験でスコアが 1 を下回っており、同化がうまくいっていないことがわかる。不連続性が最も強い 2 次の演算子を除いて微分可能な場合と不可能な場合の結果を比較すると、JH のほうがスコアの変化が小さく不連続性に対してロバストである。これは、今回の観測演算子を接線形近似したとき、微分可能な場合と不可能な場合で変わるのは符号のみであり、この符号の変化は 2 次関数形を仮定するコスト関数では打ち消されるためと考えられる。

ここで、ETKF-FH op-av が発散した理由について考察する。観測誤差 10^{-3} の 3 乗観測を与えた実験で、ETKF-FH av-op と op-av の最初のサイクルにおけるカルマンゲインの構造を Fig.5 に示す。この実験設定で ETKF-FH op-av は発散する。av-op では対角成分にシグナルが集中しており、格子点と同じ位置の観測の影響が最も大きくなっているのに対して、op-av では 11 番目の格子点付近に 4 から 12 番目の観測の影響が現れており、その大きさも av-op の 20 倍になっている。この 11 番目の格子点付近は最初のサイクルで最もアンサンブルメンバー間の差 (スプ

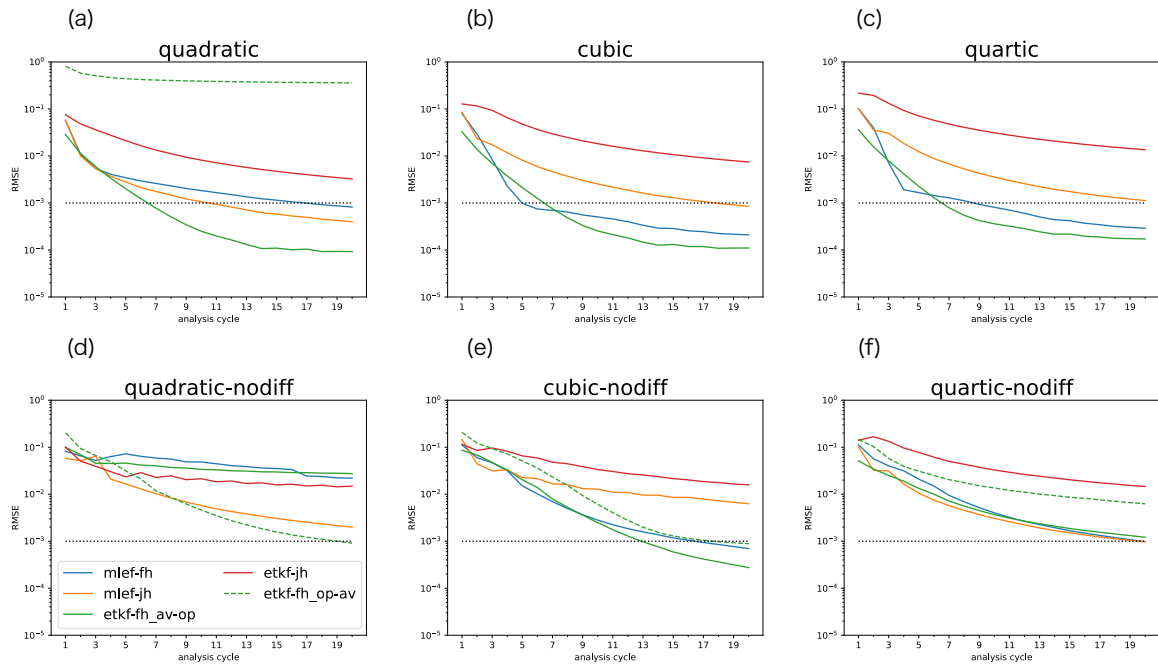


Fig. 3: Analysis RMSE of experiments with observation error of 10^{-3} . Horizontal axes are assimilation cycles, and vertical axes are RMSE by log scale. Black dot lines indicate observation error level. Legends in (d) are common to all panels.

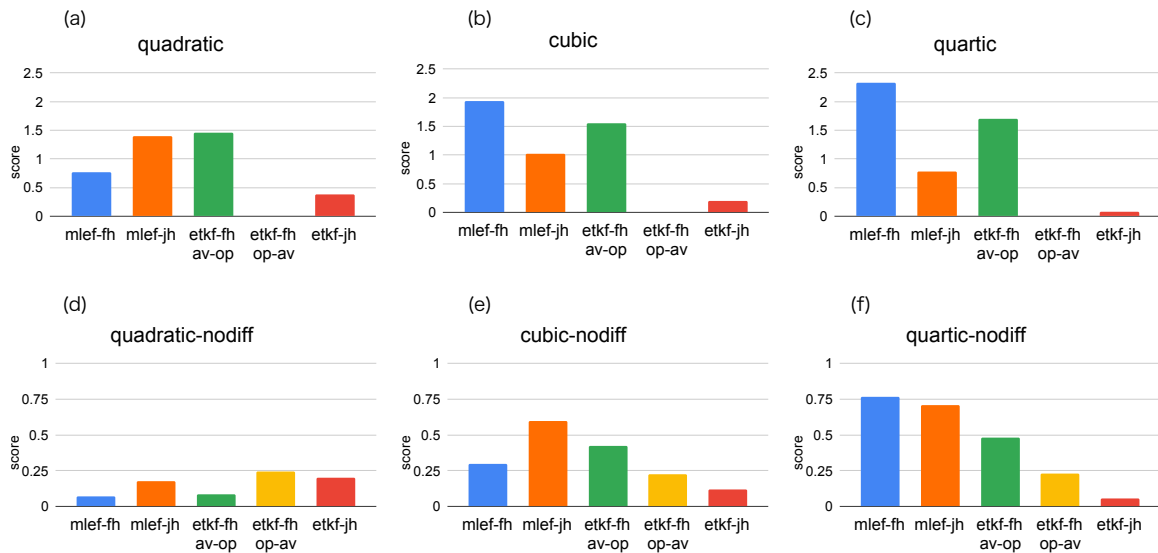


Fig. 4: Score for each observation operators. The magnitude of 1.0 means that analysis error is the same level of observation error. Note that the vertical axis ranges are different between (a,b,c) and (d,e,f).

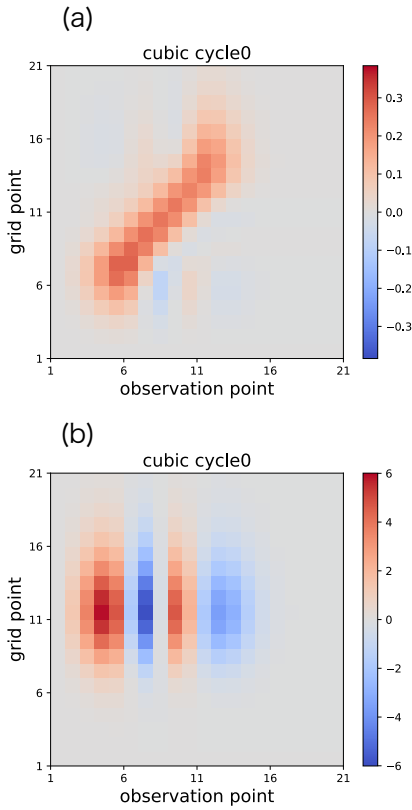


Fig. 5: Kalman gain structures for ETKF-FH (a) av-op and (b) op-av. The regions beyond first 20 points are cropped.

レッド) が大きい位置である。ここから、op-av では一つの格子点に対して広い範囲の観測が大きな影響を与えるような重みが作られ、発散の原因となったことがわかる。

このようなカルマンゲインを創り出した原因について、カルマンゲインを特異値分解を利用してモード展開することによって分析する。カルマンゲインの構成式 (24) において、 \mathbf{Z} 行列を特異値分解

$$\begin{aligned} \mathbf{Z} &= \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T \\ \mathbf{\Sigma} &= \text{diag}(\sigma_1, \dots, \sigma_m) \\ \mathbf{U} &= (\mathbf{u}_1, \dots, \mathbf{u}_m), \mathbf{u}_i \in \mathbb{R}^P \\ \mathbf{V} &= (\mathbf{v}_1, \dots, \mathbf{v}_m), \mathbf{v}_i \in \mathbb{R}^m \end{aligned} \quad (33)$$

し、(24) に代入し、 $\mathbf{R} = \sigma_o^2 \mathbf{I}_p$ (σ_o は観測誤差) であることを利用すると、

$$\begin{aligned} \mathbf{K} &= \mathbf{X}^f (\mathbf{I}_m + \mathbf{V}\mathbf{\Sigma}^2\mathbf{V}^T)^{-1} \mathbf{V}\mathbf{\Sigma}\mathbf{U}^T \mathbf{R}^{-1/2} \\ &= \mathbf{X}^f \mathbf{V} (\mathbf{I}_m + \mathbf{\Sigma}^2)^{-1} \mathbf{\Sigma}\mathbf{U}^T \mathbf{R}^{-1/2} \\ &= \sum_{i=1}^m \boxed{\frac{\sigma_i}{(1 + \sigma_i^2)\sigma_o}} \mathbf{X}^f \mathbf{v}_i \mathbf{u}_i^T \end{aligned} \quad (34)$$

となり、特異値 σ と観測誤差 σ_o によって決まるスケール (囲み部分) を重みとするモード (下線部分) の重ね合わせによって決まる。

\mathbf{Z} 行列の特異値とスケール、カルマンゲインをモード展開した結果を Fig.6 に示す。アンサンブル摂動行列 \mathbf{X}^f は列方向の平均が 0 のため、行列のランクは高々 $m - 1$ である。そのため、 \mathbf{X}^f を元に構成される \mathbf{Z} のランクも高々 $m - 1$ であることが期待される。ETKF-FH av-op (26) ではランク落ちの性質が保持されるため、特異値は 4 つしか存在しない。一方 ETKF-FH op-av (27) では平均操作の後に非線型の観測演算子を作用させるため列方向の平均が 0 とならず、従って 5 つ目の特異値が存在している。しかしこの 5 つ目の特異値は本来存在しなかったものであるため、他の 4 つの特異値に比べて大きさが小さい。式 (34) から、モード展開したときのスケールは特異値の大きさに反比例するため、この 5 つ目の特異値に対応するモードの寄与は相対的に大きくなる。そのため、5 番目の偽の特異値に対応する異常に大きい構造が物理的に不安定な解を作り出す原因であると考えられる。 \mathbf{Z} 行列のランクの保存は解析アンサンブル摂動行列 (29) のランクにも影響を与える。解析アンサンブルのランクが保たれる (列方向の平均が 0 となる) ことは、アンサンブル平均と整合する解析アンサンブルを作成する上で重要である (Hunt et al. 2007)。以上より、非線型の観測演算子にアンサンブルカルマンフィルタを適用する際は、線形の場合に満たすべき性質を保持するように拡張を行う必要がある。

4.2 接線形近似の影響

この節では MLEF に焦点を戻し、接線形近似の有無による違いについて考察する。

Fig.4 に示されるように、観測の非線形性が強くなるほど接線形近似を用いない方が精度が良くなり、これは想定される結果である。しかし、2 次の観測の場合や微分不可能な場合は、接線形近似を用いない導出の優位性が現れない。これはコスト関数の最適化の収束性に関わる問題であることが考えられる。例として、微分不可能な 2 次の観測を同化した実験の最初のサイクルにおける、最小値探索中の勾配ノルムの変化を Fig.7(a) に示す。勾配ノルムは最小値探索の収束判定に用いられる。青線で示す MLEF-FH の変化が途中で消えているが、これは最小値探索が

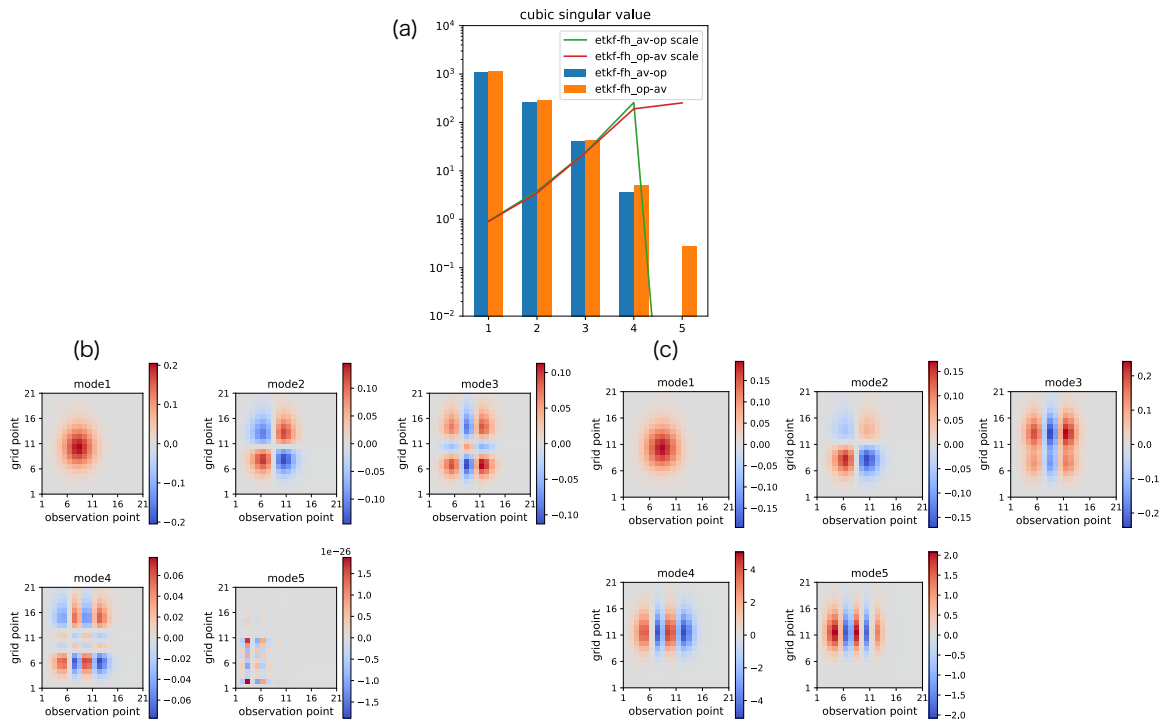


Fig. 6: (a) Singular values (bars) and scales (lines) for matrix \mathbf{Z} . The definition of scales is described in the text. (b,c) Decomposition of Kalman gain for ETKF-FH av-op (b) and op-av (c). The regions beyond first 20 points are cropped.

収束したためではなく探索が失敗していることを示している（勾配ノルムは 10^3 程度までしか下がっていない）。MLEF-FH はより正確に観測項を評価できる一方で、観測項に2次以上の高次の項を含むためコスト関数の形状が2次関数からはかけ離れたものとなる。本研究で用いた非線形共役勾配法を含む一般に用いられる最適化手法は、最小値を求める目的関数が2次関数に近いことを仮定しているため、MLEF-FH の最小値探索は収束しづらくなる。従って、MLEF-FH の解析精度は観測評価の正確性とコスト関数の最適化の収束性とのバランスによって決まると考えられる。

第2.1節で述べたように、MLEFでは最適化の収束性を高めるためにヘシアンを利用した前処理を行う。本来のアルゴリズムでは、前処理行列(13)は最適化を始める前に計算し、最適化の最中は固定する。線形の観測演算子の場合、一回の前処理でコスト関数の形状を最適値を中心とする m 次元球に修正することができるため、最適化は1回の反復で終わる。しかし非線形の場合、前処理は初期値の近傍においてコスト関数の形状を m 次元球に近づけることに留まり、予報値と解析値（最適値）が大きく離れてい

る同化初期には、一回の前処理だけでは十分に収束性を高めることができない。そこで、最適化の数回の反復ごとに前処理行列を再評価する「リスタート法」を導入する。前処理行列の再評価は最適化が失敗するか、反復が制御変数の次元 m ($= 4$) を超えた段階で行う。今回の実験ではリスタートは最大で20回までとし、リスタートを導入しても解に変化がなくなった時点で収束判定の結果によらず反復を終了するように設定した。リスタート法を導入した場合の勾配ノルムの変化を Fig.7(b) に示す。MLEF-FH はまだ完全には収束していないが、導入前と比較して勾配ノルムが小さくなり、より最適値に近づいていることがわかる。また MLEF-JH でも収束が早くなっていることがわかる。

リスタート法の導入、及び最適化手法の選択による解析精度への影響を評価するために、以下に挙げる5種類の最適化手法を用いてリスタート法の有無によるスコアの変化を調べた (Fig.8).

勾配のみ用いる手法

- 非線形共役勾配法 (Conjugate Gradient)
- メモリ制限付き BFGS 法 (Limited memory BFGS)

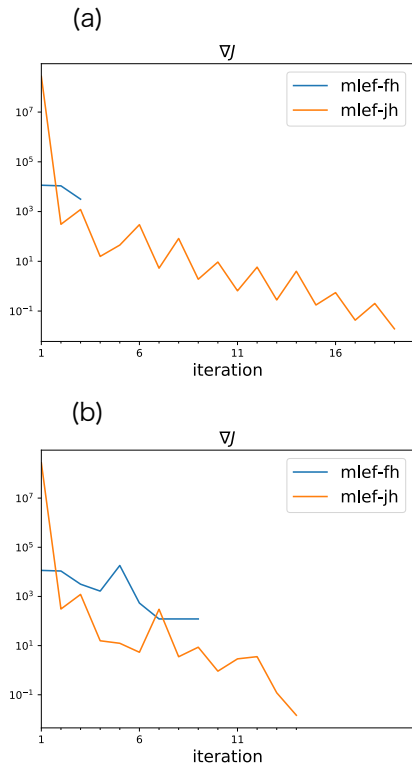


Fig. 7: Gradient norms during the first cycle minimization of the cost function with non-differentiable quadratic observations: (a) without restart method and (b) with restart method. Horizontal axes indicate iteration number.

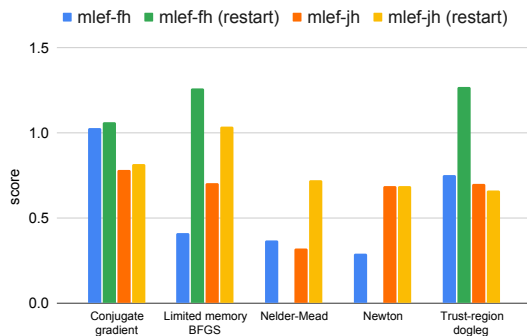


Fig. 8: Score for each minimization methods with or without restart method. Score is averaged over all observation operators.

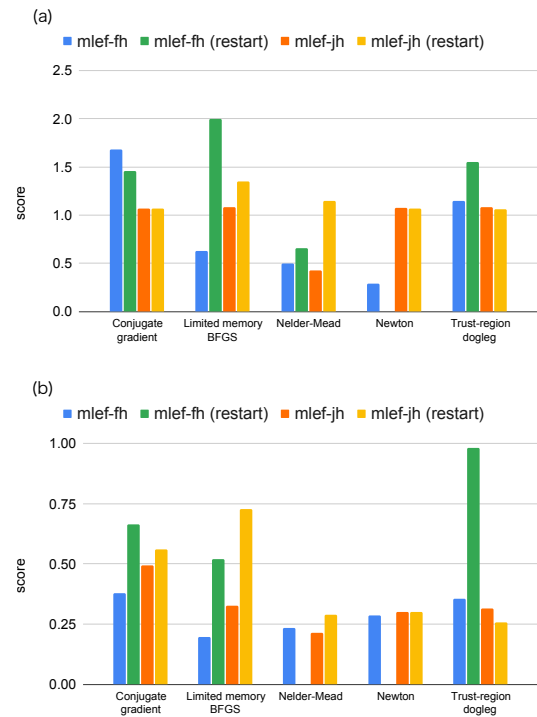


Fig. 9: As for Fig. 8, but with (a) differentiable observation operators, and (b) non-differentiable operators.

勾配を用いない手法

- Nelder-Mead 法

勾配とヘシアンを用いる手法

- Newton 法
- 信頼領域 dogleg 法 (Trust-region dogleg)

リスタート法を導入すると、FH と JH とともに多くの手法で精度が改善する。FH の方が改善率は大きい。しかし、FH の場合に Nelder-Mead 法と Newton 法でリスタート法を導入すると、計算が不安定になる。この 2 手法では既定の反復回数内で最適値に収束したり停滞したりすることなく解が変化し続けてしまい、最適値から離れた位置で止まってしまうために計算不安定に繋がっていた。手法間で比較すると、FH でメモリ制限付き BFGS 法と信頼領域 dogleg 法にリスタート法を導入した場合が最も良い精度が得られた。リスタート法はヘシアンを単位行列に近づけるため、反復の中でヘシアンの近似を更新していくメモリ制限付き BFGS 法、及び現在の位置の近傍で目的関数を 2 次関数近似する信頼領域法との相性が良いと考えられる。微分可能性で分けて比較すると (Fig.9), メモリ制限付き BFGS 法は微分可能な観測に対して、信頼領域 dogleg 法は微分不可能な観測

に対して特に精度の改善が見られた。

5. まとめ

本研究では、最尤法アンサンブルフィルタの非線型観測同化に対する性能を、1次元の移流拡散方程式を予報モデルとする同化実験でアンサンブル変換カルマンフィルタと比較することにより調査した。

最尤法アンサンブルフィルタは変分法を基礎とし、コスト関数をアンサンブル空間で最適化することによって解析値を求める。アンサンブル変換カルマンフィルタとの比較から、このコスト関数の最適化が非線型観測演算子に対して有効であることがわかった。また、アンサンブル摂動を利用して観測演算子の接線形近似を避けた同化手法を用いると、特に非線型性が強い場合により良い解析精度を示した。最尤法アンサンブルフィルタは微分不可能な観測演算子に対しても微分可能な場合と同様に同化できることが先行研究で指摘されている (Zupanski et al., 2008) が、今回の実験では微分不可能な観測に対する解析精度は観測の精度を上回ることができなかった。これは不連続を含む関数の最適化問題の難しさに起因すると考えられる。

アンサンブル変換カルマンフィルタではアンサンブル平均を解析変数とするため、接線形近似を避ける定式化が2通り存在する。観測演算子を作用させる前にアンサンブル平均をとって定式化すると、強い非線型性を持つ観測を同化した際に物理的に不安定な解が生成される。これは解析値の一部に広い範囲の観測が影響を及ぼすためであり、カルマンゲインの構造に原因がある。特異値分解による解析から、アンサンブル平均をとった後に観測演算子を作用させると、観測空間に射影したアンサンブル摂動行列のランクが維持されず本来存在しないはずの特異値が現れてしまい、この偽の特異値に対応するモードが異常な重みを生み出していたことがわかった。これはアンサンブルによってカルマンゲインを表現する全てのアンサンブルカルマンフィルタに共通する問題であり、元々線形または弱い非線形を仮定して導出されているカルマンフィルタにおいては、非線型の場合でも線形の場合に満たされる性質をなるべく保持するように拡張する必要があることを示唆している。なお、微分不可能な場合は不連続性のために、観測演算子と平均をとる操作の順番によらず観測空間のアンサンブル摂動のランクは保持されない。

しかし5番目の特異値の大きさが他の4つの特異値と同程度であるために、偽のモードの寄与が微分可能な場合ほど大きくならず発散しなかったと考えられる。

2次の観測や微分不可能な観測の場合に、最尤法アンサンブルフィルタが期待される解析精度を示さなかった。これは観測項に含まれる高次の項の影響で、コスト関数の最適化が収束しづらいためであると考えられる。最尤法アンサンブルフィルタでは収束性を高めるための前処理を導入しているが、最適化の反復開始の前に行う前処理の効果は強い非線形性を持つ観測演算子を用いる場合限定的である。そこで、反復の最中に前処理を再度行うリスタート法を導入した。このリスタート法は最適化の収束性を改善するが、完全に収束するまでは至らなかった。最適化手法間の比較から、リスタート法はヘシアンまたはその近似を利用する手法と組み合わせると最も良い精度を示すことがわかった。ただし手法によっては計算不安定を引き起こすため、リスタート法の有効性は問題依存である。今回比較した中では信頼領域 dogleg 法が微分不可能な観測に対して最も有効であり、将来の応用に向けて有望な知見が得られた。

本研究から、最尤法アンサンブルフィルタが非線形観測同化に対して有効な同化手法であることが示された。しかし本研究は現実の気象を簡略化したモデルを用いているため、実際の気象予報への応用のためにはさらなる調査が必要である。例えば今回の実験では共分散膨張や局所化なしでも十分な精度が得られたが、多変数・高次元のモデルに適用するためには、アンサンブル数の少なさによるサンプル誤差を緩和するためのこれらの手法が必要になってくると考えられる。特に局所化を最尤法アンサンブルフィルタの最適化に反映させるためには工夫が必要であり、どのような局所化が有効であるかは明らかでない。今後2次元、3次元のモデルへの適用を通して効果的な局所化手法を検討していく。さらに、モデルの誤差や観測欠損への応答について調査することも重要である。

また、微分不可能な観測の同化性能は期待されるほど高くはなかった。これは上述したように最適化の収束性が大きな課題であると考えられる。前章の最後に示した最適化手法間の比較から、信頼領域法が微分不可能な観測同化に適していることが示唆されたが、より多次元の問題になった場合には計算コ

ストと精度のバランスも重要な要素となるため、どの手法が最適であるかは明らかではない。各最適化手法の特性を理解し、最適な手法を探求していく。

参考文献

- Bishop, C. H., B. J. Etherton, and S. J. Majumdar, (2001): Adaptive sampling with the Ensemble Transform Kalman Filter. Part I: Theoretical aspects., *Mon. Wea. Rev.*, **129**, pp.420–436.
- Burgers, J. M., (1948): A mathematical model illustrating the theory of turbulence., *Adv. App. Mech.*, pp.171–199.
- Fletcher, R., and C. M. Reeves, (1964): Function minimization by conjugate gradients., *Comp. J.*, **7**, pp.149–154.
- Hunt, B. R., E. J. Kostelich, and I. Szunyogh, (2007): Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter., *Physica D*, **233**, pp.112–126.
- Zupanski, M., (2005): Maximum Likelihood Ensemble Filter: Theoretical aspects., *Mon. Wea. Rev.*, **133**, pp.1710–1726.
- Zupanski, M., I. M. Navon, and D. Zupanski, (2008): The Maximum Likelihood Ensemble Filter as a non-differentiable minimization algorithm., *Quart. J. Roy. Meteor. Soc.*, **134**, pp.1039–1050.

(論文受理日: 2021 年 8 月 31 日)