

アンサンブル学習を用いた粒子フィルタの提案

Particle Filter using An Ensemble Learning

山田和明*¹
Kazuaki YAMADA*¹東洋大学
Toyo University

This paper proposes a new particle filter using the ensemble learning. The proposed method, firstly, generates some particle groups that have different initial hyper parameters, by uniform random numbers, and secondly, operates a particle filter for each particle group, and finally, estimates a system state by integrating the estimated values of some particle groups by using an ensemble learning. The proposed method uses both a simple average and a weighted average that uses the logarithm likelihood. We investigate by estimating the system states, system noise and observation noise in the small nonlinear system and observation model.

1. はじめに

粒子フィルタは、あるシステムの時刻 t における状態 x_t を観測値 y_t からモンテカルロ法を用いて推定する時系列フィルタの一種である。実装が比較的簡単でシステムの状態を精度良く推定できるため、気象学・海洋学、経済学、ロボットの位置推定や画像処理など幅広く利用されている。

粒子フィルタは、式 (1-a) の時刻 t におけるシステムの状態 x_t を表すシステムモデルと、式 (1-b) のシステムの状態とセンサによる観測値の関係を表す観測モデル、システムとセンサに発生するシステムノイズ v_t と観測ノイズ w_t が既知の場合、各ノイズの大きさに応じて適切な重み付けを行い、時刻 t におけるシステムの状態 x_t を精度良く推定することができる。

$$x_t = f(x_{t-1}, v_t), \quad v_t \sim N(0, \sigma^2) \quad (1-a)$$

$$y_t = h(x_t, w_t), \quad w_t \sim N(0, \tau^2) \quad (1-b)$$

粒子フィルタにおけるシステムノイズはシステムをモデル化する際に生じるズレに起因するものであり、観測ノイズは観測時に発生するノイズである。粒子フィルタでは、これらのパラメータを超パラメータと呼ぶ。これらの超パラメータがどのような問題においても常に正確に測定できるとは限らないため、粒子フィルタの適用範囲を拡大するためには、超パラメータをシステムの状態と同時に推定する必要がある。北川は超パラメータをシステムの状態に含めて同時推定する自己組織化状態空間 (Self-Organization State Space : SOSS) モデル [1] を提案している。しかし、超パラメータに関する事前知識が乏しい場合には、自己組織化状態空間モデルにおける超パラメータの初期値分布を適切に設定することが困難であることが指摘されている [2]。そのため、矢野らは Nelder-Mead 法を用いて超パラメータの初期分布を効率的に探索するアルゴリズムを提案している。このアルゴリズムは、超パラメータを精度よく推定することが目的であるため、予め収集したデータ集合に対して繰り返し実行する非逐次型の推定手法である。

本論文では、超パラメータに関する事前知識が乏しいとき、逐次的に超パラメータを推定しながら推定精度を向上させるために、天気予報などに用いられるアンサンブル学習 [3] を応

用した粒子フィルタ手法を提案する。アンサンブル学習は、アンサンブルごとに異なる初期状態から推定を始め、各アンサンブルの平均を取ることでシステムの状態を推定する。その際、各アンサンブルのばらつきを示すスプレッドから推定結果がどの程度信頼できるかを判断することができる。提案手法は、超パラメータの異なる初期粒子集団を複数生成し、各粒子集団に対して粒子フィルタを適用し、各粒子集団の推定値の加重平均をとることでシステムの状態を推定する。本稿では、各粒子集団の対数尤度を加重として用いる場合と、単純平均を用いる場合の推定精度を比較する。また、粒子フィルタの推定精度とスプレッドの関係について分析する。

本稿は、次章において提案手法であるアンサンブル学習を用いた粒子フィルタについて説明し、3章において提案手法の有効性を検証するため、テスト問題に適用する。最後にまとめと今後の課題について述べる。

2. アンサンブル学習を用いた粒子フィルタ

本章では、まず、粒子フィルタの推定方法について説明し、次にアンサンブル学習を用いた粒子フィルタの概要を説明する。そして、アルゴリズムについて詳述する。

2.1 提案手法の概要

粒子フィルタの基本的な考え方は、図 1 に示すように、まず、多数の粒子をシステムモデルに基づいて状態空間に散布し、システムの状態を近似する予測分布を生成する。次に、観測値 y_t に基づいて各粒子の値がシステムの状態 x_t にどれだけ当てはまるかを尤度により評価する。そして、尤度に基づいて粒子をリサンプリング (復元抽出) し、フィルタ分布を生成する、という操作を繰り返すことでシステムの状態を推定する。

提案手法は、図 2 に示すように初期パラメータの異なる複数のフィルタ分布 (粒子集団) で状態空間を覆い、各粒子集団に対して粒子フィルタを実行し、各粒子集団の推定値の加重平均をとることでシステムの状態を推定する。また、提案手法に自己組織化状態空間モデルを用いる場合は、図 2(b) に示すようにシステムモデルと観測モデルのノイズの分散 (超パラメータ) からなる超パラメータ空間を分割し、分割した超パラメータ空間に初期値が異なる粒子集団を一様乱数により生成する。このように提案手法は、各粒子集団に対して粒子フィルタによりシステムの状態を推定し、その結果をアンサンブル学習に

連絡先: 山田和明, 東洋大学理工学部機械工学科, 〒 350-8585
埼玉県川越市鯨井 2100, yamadak@toyo.jp

より統合するため、粒子フィルタで用いられる自己組織化状態空間モデルや固定ラグ平滑化などの手法をそのまま利用することができる。そのため本稿では、粒子フィルタ (Particle Filter : PF) および固定ラグ平滑化を導入した粒子スムージング (Particle Smoothing : PS) にアンサンブル学習を導入した場合の推定精度を比較する。なお、本稿では、アンサンブル学習として、各粒子集団の単純平均を用いる場合と、各粒子集団の対数尤度を加重として用いる場合の 2 種類を試す。

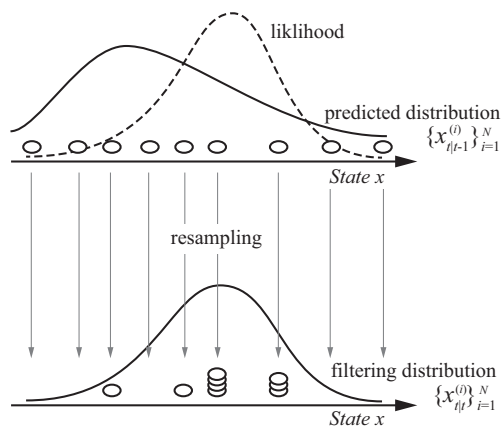


図 1: A conceptual model of a particle filter.

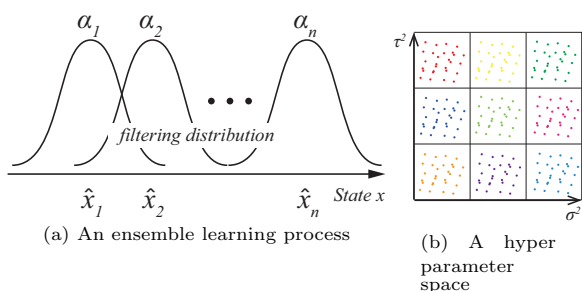


図 2: A conceptual diagram of a particle filter using an ensemble learning.

2.2 アルゴリズム

提案手法のアルゴリズムは以下のように記述できる。

- (1) システムモデルと観測モデルのノイズの分散の推定範囲をそれぞれ n , m 等分し、合計 $n \times m$ 個の粒子集団を生成する。
- (2) 各粒子集団において、初期分布を近似する粒子集団 $\{x_{0|0}^{(i)}\}_{i=1}^N$ ($x_{0|0}^{(i)} \sim p_0(x)$) を一様乱数により散布する。ただし、 $p_0(x)$ は x の時刻 $t = 0$ における初期分布を表す。
- (3) 各粒子集団において、 $t = 1, \dots, T$ について (a), (b), (c) のステップを実行する。

(a) 尤度計算

各 i ($i = 1, \dots, N$) について (i)~(iii) を実行する。

- (i) 乱数 $v_t^{(i)} \sim q(v)$ を生成する。

- (ii) $x_{t|t-1}^{(i)} = f_t(x_{t-1|t-1}^{(i)}, v_t^{(i)})$ を計算する。

- (iii) $\beta_t^{(i)} = p(y_t | x_{t|t-1}^{(i)})$ を計算する (式 (2))。

(b) リサンプリング

粒子集団 $\{x_{t|t-1}^{(i)}\}_{i=1}^N$ から $\tilde{\beta}_t^{(i)} = \beta_t^{(i)} / \sum_{i=1}^N \beta_t^{(i)}$ の確率で $x_{t|t-1}^{(i)}$ を重複を許して抽出し、新たな粒子集団 $\{x_{t|t-1}^{(i)}\}_{i=1}^N$ を生成する。

(c) 時刻 t の状態推定

$$\hat{x}_{t|t} = \frac{1}{N} \sum_{i=1}^N x_{t|t}^{(i)}$$

- (4) 各粒子集団の状態推定値の加重平均を式 (3) より計算する。なお、本提案手法では加重 α_j を各粒子集団の対数尤度 l_j から求めている。

なお、観測ノイズ w_t が平均ゼロ、分散 τ^2 の正規分布に従う場合、粒子 $x_{t|t-1}^{(i)}$ の尤度 $p(y_t | x_{t|t-1}^{(i)})$ は次式で求められる。

$$p(y_t | x_{t|t-1}^{(i)}) = \frac{1}{\sqrt{2\pi\tau^2}} \exp \left[-\frac{\left(y_t - H(x_{t|t-1}^{(i)}) \right)^2}{2\tau^2} \right] \quad (2)$$

$$\hat{x}_t = \frac{\sum_{j=1}^{n \times m} \alpha_j \hat{x}_j}{\sum_{k=1}^{n \times m} \alpha_k} \quad (3-a)$$

$$\alpha_j = \frac{\log(l_j)}{\sum_{k=1}^{n \times m} \log(l_k)} \quad (3-b)$$

$$l_j = \log \left(\sum_{i=1}^N p(y_t | x_{t|t-1}^{(i)}) \right) - \log(N) \quad (3-c)$$

2.3 自己組織型状態空間モデル

自己組織型状態空間モデルは、式 (4) に示すように、状態 x_t に超パラメータ λ_t を含めた状態ベクトル z_t を生成する。ただし、超パラメータ λ_t はシステムノイズ v_t の分散 σ_t^2 と観測ノイズ w_t の分散 τ_t^2 から構成されている。生成した状態ベクトル z_t に対して粒子フィルタと同様の操作を繰り返すことで、システムノイズと観測ノイズの分散を同時に推定することができる^{*1}。このとき、システムモデルと観測モデルは式 (5) と記述でき、非線形関数 F と H は式 (6) となる。

$$z_t = \begin{bmatrix} x_t \\ \lambda_t \end{bmatrix}, \quad \lambda_t = \begin{bmatrix} \log \sigma_t^2 \\ \log \tau_t^2 \end{bmatrix} \quad (4)$$

$$z_t = F(z_{t-1}, v_t) \quad (5-a)$$

$$y_t = H(z_t, w_t) \quad (5-b)$$

$$F(z_{t-1}, v_t) = \begin{bmatrix} f(x_{t-1}, v_t) \\ \lambda_{t-1} + \epsilon_t \end{bmatrix} \quad (6-a)$$

$$H(z_t, w_t) = h(z_t, w_t) \quad (6-b)$$

ただし、式 (6-a) の $\lambda_t = \lambda_{t-1} + \epsilon_t$ は、超パラメータ λ_t の時間変化を表し、 $\epsilon_t = [\zeta_t, \eta_t]'$ はそれぞれ $\zeta_t \sim N(0, \nu^2)$, $\eta_t \sim N(0, \xi^2)$ とする。 ν と ξ は超パラメータを特徴付けるパラメータであるため、超々パラメータと呼ばれる。

*1 σ^2 と τ^2 は正值性を保つため対数値を用いる。

2.4 固定ラグ平滑化

固定ラグ平滑化 [5] は、状態ベクトル $z_{t|t-1}^{(i)}$ を式 (7) のように拡張し、この拡張状態ベクトルに対して粒子フィルタと同様にリサンプリングすることで実現できる。最後にリサンプリングにより得られた粒子集団 $\{z_{i|t}^{(i)}\}_{i=1}^N = \{[z_{t|t}^{(i)}, \dots, z_{t-L|t}^{(i)}]'\}_{i=1}^N$ から $\{z_{t-L|t}^{(i)}\}_{i=1}^N$ の部分を取り出すことで、時刻 $t-L$ から時刻 t の情報に基づいて平滑化した時刻 $t-L$ の状態を推定することができる。

$$\tilde{z}_{t|t-1}^{(i)} = [z_{t|t-1}^{(i)}, z_{t-1|t-1}^{(i)}, \dots, z_{t-L|t-1}^{(i)}]' \quad (7)$$

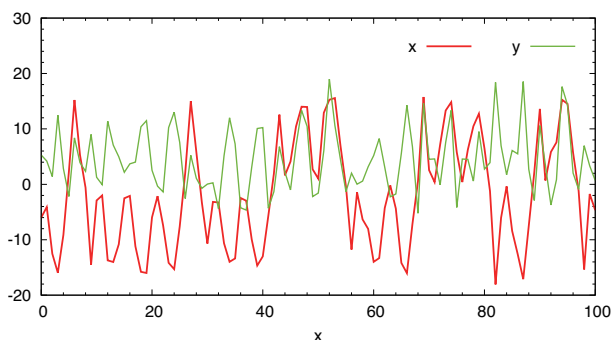


図 3: A test function.

3. 計算機実験

本章では、提案手法の有効性を検証するために非線形システム・非線形観測の小規模なシステムに対する状態推定および超パラメータの推定を行う。そして、粒子フィルタ (Particle Filter : PF) および粒子スムージング (Particle Smoothing : PS) と、これらにアンサンブル平均とアンサンブル加重平均を適用した場合の結果を比較する。

3.1 実験設定

本実験では、非線形状態の推定問題としてよく用いられる式 (8) により観測値 y_t を発生させる [4]。

$$\begin{cases} x_t = \frac{1}{2}x_{t-1} + \frac{25x_{t-1}}{1+x_{t-1}^2} + 8 \cos(1.2t) + v_t \\ y_t = \frac{x_t^2}{20} + w_t \\ x_0 \sim N(0, 5), v_t \sim N(0, 1.5), w_t \sim N(0, 8) \end{cases} \quad (8)$$

この推定問題におけるシステムの状態 x_t と観測値 y_t との関係の一例を図 3 に示す。グラフの横軸は時間を表し、赤と緑の線はそれぞれシステムの状態 x_t と観測値 y_t を表す。

実験では、自己組織化状態空間モデルの超パラメータを $\zeta = \eta = 0.05$ とし、また、アンサンブル学習を行うために、システムモデルと観測モデルのノイズの分散の範囲をそれぞれ $\sigma^2 = (0, 2)$, $\tau^2 = (0, 20)$ とし、2次元の超パラメータ空間を構成する。そして、各次元を 3×3 等分し、粒子を式 (9) に従って一様乱数により散布する。

$$\sigma_i^2 \sim \text{Uniform}(M_i^\sigma - w_i^\sigma, M_i^\sigma + w_i^\sigma) \quad (9-a)$$

$$\tau_j^2 \sim \text{Uniform}(M_j^\tau - w_j^\tau, M_j^\tau + w_j^\tau) \quad (9-b)$$

ただし、 M_i^σ と w_i^σ はそれぞれ i 分割目の中心と幅を表し、 M_j^τ と w_j^τ はそれぞれ j 分割目の中心と幅を表す。

アンサンブル学習を適用するときは各粒子集団の粒子数を 100 とし、アンサンブル学習を適用しないときは粒子数を 900 個とする。これは、アンサンブル学習を適用するとき粒子集団は $3 \times 3 = 9$ 個のため全体では 900 個の粒子を用いており、粒子数による推定精度の影響を低減するためである。

推定精度を比較するために、100 ステップを 1 トライアルとして 100 トライアル行い、トライアルごとにシステムの状態 x_t と粒子フィルタによる推定値 $x_{t|t}$ との誤差の二乗和 $\sum_n (x_t - x_{t|t})^2$ を計算し、100 トライアルの平均を比較する。なお、粒子スムージング (PS) のラグ幅は $L = 20$ とする。

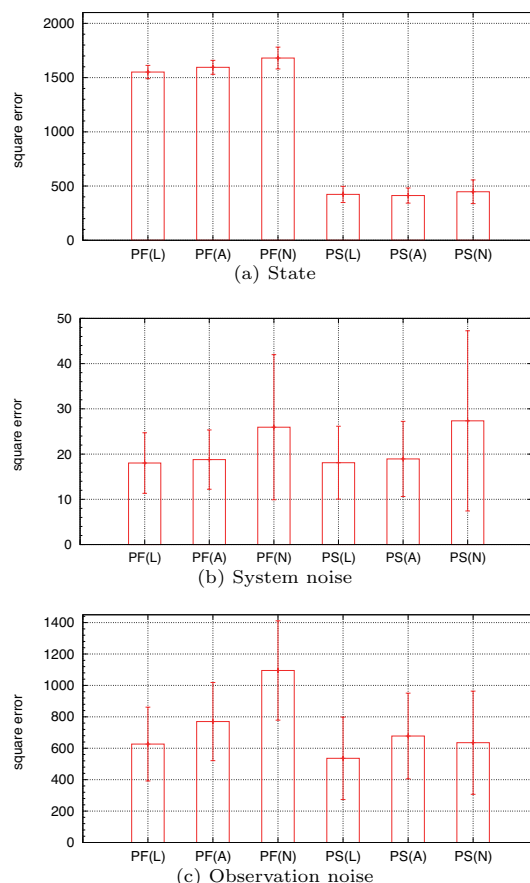


図 4: Square errors of estimated values.

3.2 実験結果

図 5 にシステムの状態、システムモデルおよび観測モデルのノイズの分散と、推定結果との誤差の二乗和の平均を示す。グラフの横軸は各推定手法を表し、縦軸は誤差の二乗和の平均を表し、標準偏差を表示している。なお、グラフの PF と PS はそれぞれ粒子フィルタと粒子スムージングを表し N, A, L はそれぞれアンサンブル学習を用いない場合、アンサンブル平均を用いた場合、アンサンブル加重平均を用いた場合を表す。例えば、PF(L) は粒子フィルタにアンサンブル加重平均を用いた場合を意味する。

図 4(a) はシステムの状態を推定したときの誤差の二乗和の平均を表す。グラフを見ると PF より PS の方が誤差が小さい、また、PF と PS にアンサンブル学習を用いた方が誤差が小さいことがわかる。同様に、システムノイズと観測ノイズの誤差の二乗和をプロットした図 4(b) と図 4(c) を見ると、アンサン

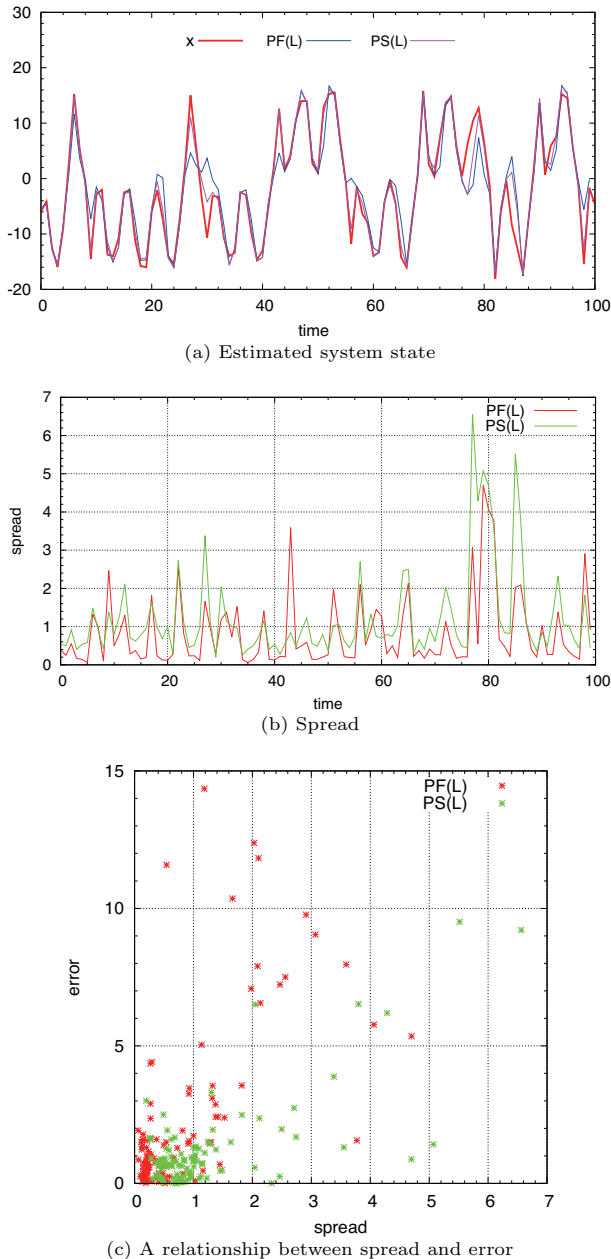


図 5: Square errors of estimated values.

ブル学習を用いた方が誤差の二乗和が小さい。しかし、図 4(c)を見ると、PS のみの場合とアンサンブル学習を用いた場合では誤差の二乗和に余り差が見られない。これは PS の平滑化により推定精度が向上したものと考えられる。以上の結果から、アンサンブル学習を用いることで、システムモデルと観測モデルのノイズの分散の推定精度が向上し、システムの状態の推定精度が改善したものと考えられる。しかし、PS の場合は、観測モデルのノイズの分散の推定精度がアンサンブル学習の有無に関わらずほぼ同じであるため、システムの状態を推定した結果に余り差が生じなかったものと考えられる。

次に、PF と PS のみ用いた場合より推定精度が良かったアンサンブル加重平均を用いた PF と PS の推定結果を図 5(a) に示す。PF と PS の推定結果のズレが大きい $t = 30$ および $t = 44$, $t = 80$ 付近において、PS の方がズレの幅が小さいこ

とから推定精度が高いことがわかる。図 5(b) はアンサンブル学習を行った際のスプレッドを表している。スプレッドはアンサンブルの分散を表し、スプレッドが大きい場合は信頼性が低いとされている [4]。図 5(b) において、推定値とのズレが大きかった $t = 30$ および $t = 44$, $t = 80$ 付近を見るとスプレッドが高いことがわかる。図 5(c) は、スプレッドと誤差の関係をプロットしたものである。PF の場合はスプレッドが 1 以下でも誤差が大きい場合が見られるが、PS の場合はスプレッドが 1 以下のとき大きな誤差は見られない。以上のことからアンサンブル加重平均を用いることで、スプレッドの値から推定結果の信頼性を推測できると考えられる。

4. おわりに

本稿では、アンサンブル学習を用いた粒子フィルタを提案した。提案手法は、超パラメータの初期値が異なる粒子集団を一様乱数により複数生成し、各粒子集団に対して粒子フィルタを実行し、各粒子集団の推定結果をアンサンブル学習により統合することで、システムの状態を推定する。提案手法では、アンサンブル学習として単純平均を用いる場合と、各粒子集団の対数尤度を加重平均に用いる場合の 2 種類の方法を用いた。

本稿では、提案手法の有効性を検証するために、非線形システム・非線形観測の小規模なシステムの状態および超パラメータの推定を行った。その結果、提案手法は、粒子フィルタのみの場合よりシステムの状態、システムノイズ、観測ノイズの推定誤差の二乗和が減少したことを確認した。また、各粒子集団のばらつきを表すスプレッドと粒子フィルタの推定誤差との間に相関関係が見られたことから、スプレッドの大きさから推定結果の信頼度を推測できる可能性があることを示した。

今後の課題として、提案手法の分析を進めてアンサンブルの計算方法の改良を行い、様々なテスト問題を通して提案手法の適用範囲を明らかにする予定である。

謝辞

本研究の一部は、JSPS 科研費 25730185, および、JST RIS-TEX 問題解決型サービス科学研究開発プログラム採択プロジェクト「価値創成クラスモデルによるサービスシステムの類型化とメカニズム設計理論の構築」の助成を受けたものです。

参考文献

- [1] 北川源四郎, 樋口知之, 知識発見と自己組織型の統計モデル, bit 別冊「発見科学とデータマイニング」, pp.159-168, 2000.
- [2] 矢野浩一, 佐藤整尚, 初期分布探索付き自己組織化状態空間モデルによる金融時系列解析の最前線:t 分布付き確率的ポラティリティ変動モデルへの応用, FSA リサーチ・レビュー 2006, pp.143-166, 2006.
- [3] 山口宗彦, 気象庁台風アンサンブル予報システム, 天気, Vol.55, No.6, pp.521-524, (2008).
- [4] 中村和幸, 上野玄太, 樋口知之, データ同化: その概念と計算アルゴリズム, 統計数理, Vol.53, No.2, pp.211-229, (2005).
- [5] 樋口知之 他, データ同化入門 (予測と発見の科学), 朝倉書店, (2011).