

ベイズ推定を用いた低次元モデルによる 動的マルチモデル EnKF 同化実験

大塚成徳・三好建正（理研・計算科学）

1 はじめに

数値天気予報において、予報の不確実性を表現する方法としてアンサンブル予報が用いられてきた。近年では、予報モデル自体の不確実性を扱うために、マルチモデルアンサンブル予報が研究されている。多くの場合、異なる力学コア、物理スキームを用いたモデルを組み合わせることにより、マルチモデルアンサンブルを構成している (e.g., Stensrud et al., 2000)。マルチモデルアンサンブルを用いたデータ同化についても研究が進められている (e.g., Meng and Zhang, 2007)。

モデルアンサンブルと初期値アンサンブルを組み合わせる場合、先行研究においては多くの場合、各モデルのアンサンブルサイズは事前に定められている。たいていの場合、各モデルに同じアンサンブルメンバ数が割り当てられている。しかし、全てのモデルに同じメンバ数を割り当てることが最適とは限らない。そこで、本研究の目的は、マルチモデルアンサンブルデータ同化における各モデルの最適なアンサンブル数を動的に推定することである。

2 手法

ここではベイズの定理に従い、 N_{model} 個のモデルに割り当てるアンサンブルメンバ数を示す離散確率分布 $P(A_i)$ ($i = 1, \dots, N_{\text{model}}$) を、観測に基づいて更新する (図 1)。

$$P(A_i|X) = \frac{P(X|A_i)P(A_i)}{P(X)}, \quad (1)$$

$$\sum_{i=1}^{N_{\text{model}}} P(A_i) = 1, \quad (2)$$

ここで $P(A_i|X)$ は更新された分布、 $P(X|A_i)$ は観測 X に基づく各モデルの尤度、 $P(X)$ は規格化係数である。

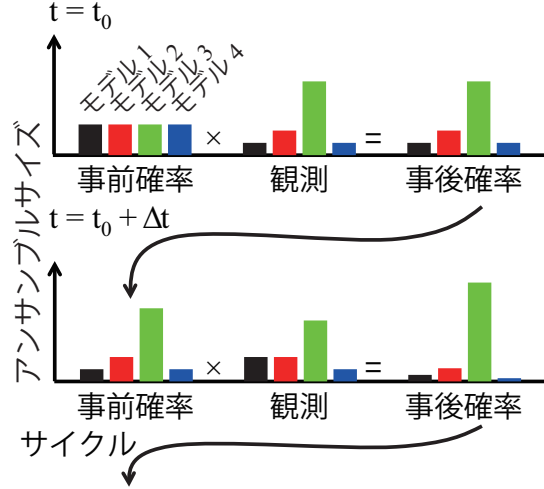


図 1: ベイズの定理に基づくモデル確率分布の更新の概念図。

観測に基づく尤度は以下のようにして決定する。まず、少し前の時刻の解析アンサンブルから延長予報を行う。次に、観測の得られた時刻において、各モデル別のアンサンブル平均と観測の距離を求め、その逆数に比例する形で尤度を与える。その尤度を用いて、観測が得られた時刻の解析後にモデル確率分布 $P(A_i)$ をベイズ更新する。図 2 に尤度決定の概念図を示す。

モデルの確率分布 $P(A_i)$ が単一の不完全モデルに収束すると、成績が低下する。そのため、モデル確率分布 $P(A_i)$ の膨張を行い、単一の不完全モデルへの収束を防ぐ。

$$\alpha = (\Delta - \Delta_0) \times \beta, \quad (3)$$

$$P(A_i) \leftarrow P(A_i) \times (1 - \alpha) + \alpha / N_{\text{model}}, \quad (4)$$

ここで、 Δ はアンサンブルカルマンフィルタの適応膨張パラメータ、 Δ_0 は完全モデル実験で得られるアンサンブルカルマンフィルタの適応膨張パラメータの平均値、 β はチューニングパラメータである。

システムの安定的な動作のため、モデル確率

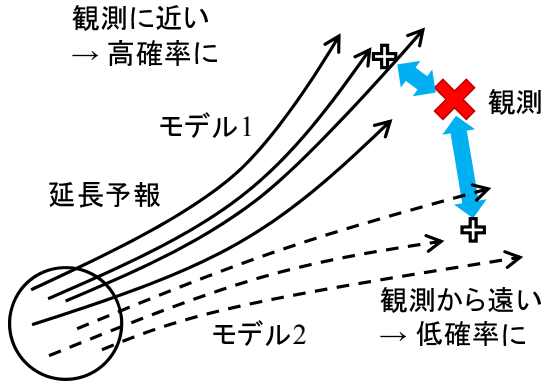


図 2: 各モデルの尤度決定方法の概念図。黒矢印（実線・破線）は位相空間上におけるマルチモデルアンサンブルの時間発展を示す。十字印は各モデルのアンサンブル平均を示す。

分布 $P(A_i)$ は時間方向に平滑化し、その分布を用いてアンサンブルサイズの変更を行う。変更の際は、アンサンブルサイズを減少させるモデルの各メンバのうち、観測にもっとも近いメンバの状態ベクトルを、アンサンブルサイズを増加させるモデルの新しいメンバの初期状態ベクトルとして用いる。

3 実験設定

使用するモデルは Lorenz の 40 変数モデル (Lorenz, 1996; Lorenz and Emanuel, 1998) である。時間積分は 4 次の Runge-Kutta を用い、時間刻みは 0.01 単位時間とする。

異なる F の値を用いたモデルを異なるモデルとみなして、マルチモデル実験を行う。本実験の目的は F の値を推定することではなく、最適なモデル組み合わせを探すことに注意されたい。以後、 $F = f_1$ のモデルを Ff_1 と表す。実験ごとに真の F を定め、真の F を用いて作成した時系列を真値、真値に分散 1 の独立な正規乱数を加えたものを観測とする。

同化手法としては逐次アンサンブル平方根フィルタ (SEnSRF, Whitaker and Hamill, 2002) を用いる。アンサンブルメンバー数は合計で 20 とし、これを複数のモデルに分配する。観測は 0.05 単位時間 (F8 の場合、6 時間相当) おきに 40 点

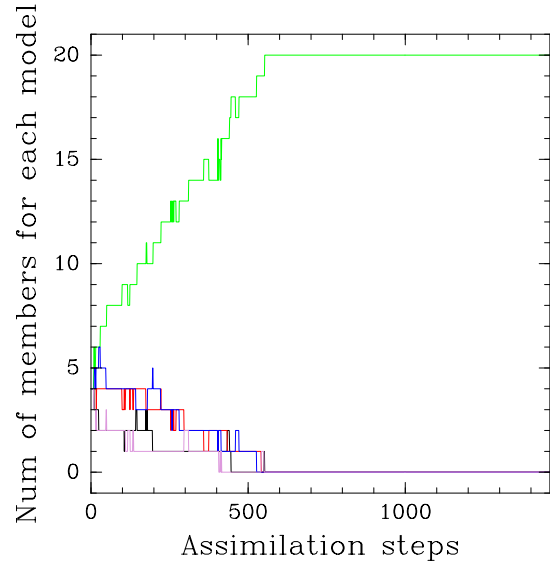


図 3: EXP1 における各モデルのアンサンブルサイズ時系列。F6:黒、F7:赤、F8:緑、F9:青、F10:紫。計算開始から 1 年を示す。

の全てで与える。実験期間は 110 年 (160600 同化ステップ) である。最初の 10 年はスピンアップとして除外し、後半 100 年を用いて統計量を計算する。誤差共分散行列の局所化スケールは 6.0 (正規分布の標準偏差を 1.0 とする単位) とした。適応共分散膨張法は Miyoshi (2011) を基に、全点で同一の膨張率 Δ を用いるようにしている。初期値は $\Delta = 1.1$ 、背景分散として 0.05^2 を用いている。

本手法において、延長予報の長さは 0.4 単位時間 (2 日相当) とした。完全モデル実験 (F8、20 メンバー、観測点 40 点) による適応共分散膨張の係数は $\Delta_0 = 1.015$ であった。モデル確率分布の $P(A_i)$ 膨張パラメータ β と、 $P(A_i)$ の時間平滑化の強さは最適にチューニングした。

4 結果

4.1 EXP1: 真のモデルを含む場合

まず、真の F を 8 とし、F6, F7, F8, F9, F10 でマルチモデルアンサンブルを構成した場合の結果を示す。この場合、真のモデル F8 が含まれ

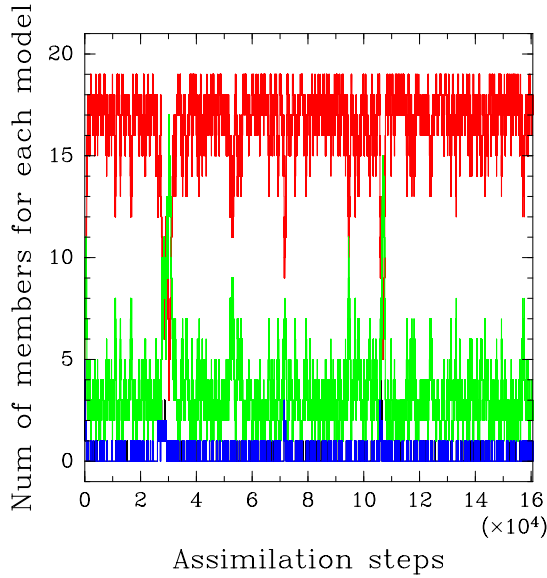


図 4: EXP2 における各モデルのアンサンブルサイズ時系列。F6:黒、F7:赤、F9:緑、F10:青。110 年分の時系列を示す。

るため、真のモデルのみに収束することが期待される。図 3 に各モデルのアンサンブルサイズの時系列を示す。同化サイクル開始と同時に真のモデル F8 のアンサンブルサイズが増加し、すべてのアンサンブルメンバが F8 に割り当てられた状態で落ち着く。すなわち、本手法は期待通りの動作をしているといえる。解析値の root mean square error (RMSE) は 0.188 で、真のモデルのみで同化を行った場合とほぼ同じになる。

4.2 EXP2：真のモデルを含まない場合

次に、真の F を 8 とし、F6, F7, F9, F10 でマルチモデルアンサンブルを構成した場合の結果を示す。この場合、真のモデル F8 は含まれないため、複数のモデルの組み合わせに落ち着くことが期待される。図 4 に各モデルのアンサンブルサイズの時系列を示す。同化サイクル開始とともに真のモデルから離れている F6, F10 が減少し、F7 と F9 が残る。F7 と F9 の割合は平均するとおよそ 8 割強と 2 割弱である。これは手動で求めた最適なモデル組み合わせとほぼ一致する。そのため、解析値の RMSE も最適なモ

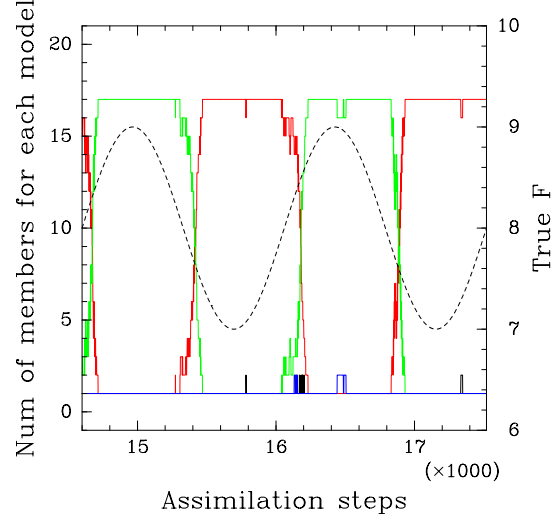


図 5: EXP3 における各モデルのアンサンブルサイズ時系列。F6:黒、F7:赤、F9:緑、F10:青。破線は真の F を示す。計算開始後 10 年から 12 年までの 2 年を示す。

デル組み合わせの場合とほぼ同じ 0.274 になる。これは、すべてのモデルに同じメンバ数を割り当てた場合の 0.316 に比べて 1 割以上 RMSE を減少させたことになる。

4.3 EXP3：真の F が時間変化する場合

最後に、真の F を $8 + \sin(2\pi t/73)$ とした場合（1 年周期相当の変動）に、F6, F7, F9, F10 で構成した動的マルチモデルアンサンブルが F の時間変化に追従できるかどうかを調べる。この実験では、各モデルに最低 1 メンバを割り当てることで、あるモデルから別のモデルへの遷移をなめらかに行うことができるようにする。図 5 に各モデルのアンサンブルサイズの時系列を示す。真の F が 7 に近づくと F7 のメンバ数が増え、9 に近づくと F9 のメンバ数が増える。真の F の位相に対して、アンサンブルサイズの位相はやや遅れる。F6 と F10 は 1 メンバ（最低メンバ数）のまま、ほとんど変化しない。

提案手法による解析の RMSE は 0.273 であり、時間的に一定な最適モデル組み合わせにおける 0.287 よりもさらに低くなる。

5 まとめ

マルチモデルアンサンブルカルマンフィルタにおいて、各モデルのアンサンブルサイズを動的に最適化するために、離散ベイズフィルタを用いた手法を開発した。解析アンサンブルからの延長予報を用いて、各モデルのアンサンブル平均と観測の距離に逆比例する形で各モデルの尤度を決定した。また、単一の不完全モデルに収束しないように、モデル確率分布の膨張を行った。

Lorenz の 40 変数モデルを用いた実験により、提案手法が良好に動作することを確認した。ここではモデル中のパラメータ F を変えたモデルを複数用いることにより、マルチモデルアンサンブルを模した実験を行った。真のモデルを含むマルチモデルアンサンブルの場合には真のモデルに収束し、真のモデルを含まない場合には複数のモデルの最適な組み合わせを得た。解析値の RMSE は、手動で求めた最適なモデル組み合わせの場合とほぼ同じ値を得た。

真の系が時間変動する場合には、その時間変動に追従する形で最適な組み合わせを得た。この場合、時間的に一定なモデル組み合わせにおける最適値と比べてさらに RMSE を減少させることができた。

参考文献

- Lorenz, E. N., 1996: Predictability: a problem partly solved. *Proc. Seminar on Predictability*, ECMWF, Reading, Berkshire, UK, Vol. 1, 1–18.
- Lorenz, E. N. and K. A. Emanuel, 1998: Optimal Sites for Supplementary Weather Observations: Simulation with a Small Model. *J. Atmos. Sci.*, **55**, 399–414.
- Meng, Z. and F. Zhang, 2007: Tests of an ensemble Kalman filter for mesoscale and regional-scale data assimilation. Part II: Imperfect model experiments. *Mon. Wea. Rev.*, **135**, 1403–1423.

Miyoshi, T., 2011: The gaussian approach to adaptive covariance inflation and its implementation with the local ensemble transform Kalman filter. *Mon. Wea. Rev.*, **139**, 1519–1535.

Stensrud, D. J., J.-W. Bao, and T. T. Warner, 2000: Using initial condition and model physics perturbations in short-range ensemble simulations of mesoscale convective systems. *Mon. Wea. Rev.*, **128**, 2077–2107.

Whitaker, J. S. and T. M. Hamill, 2002: Ensemble data assimilation without perturbed observations. *Mon. Wea. Rev.*, **130**, 1913–1924.