

アンサンブル予報におけるモデル誤差の影響 ~ EnKF を用いて ~

小山博司 (北大・環境科学)・渡部雅浩 (東大・気候システム)

1 はじめに

大気の数値予報において初期値が不正確であることによる予報の悪化を防ぐために初期値アンサンブル法が用いられ、現在までにさまざまな手法が考案されている。その内の1つであるアンサンブル・カルマンフィルタ (Ensemble Kalman Filter: EnKF, Evensen 1994) は、データ同化と初期値アンサンブル予報を融合した手法であり、解析値と高度な初期摂動の作成を同時に効率的に行う。近年では大気大循環モデルや実際の観測データへの適用など実用化へ向けた研究が行われている。

一方、数値予報モデルの不完全であることによる予報の悪化を防ぐためにモデルアンサンブル法が用いられる。その手法は一般に不完全さを持つ複数のモデル、パラメタリゼーション、パラメータなどによる予報を平均することで予報を改善させるものであり、初期値アンサンブル法に比べて有効性に理論的な根拠を持つような方法が確立していない。モデルアンサンブル法の一つであるマルチパラメータ法はパラメタリゼーションにおいて異なる複数のパラメータ値を用いる方法で、同じモデルかつ同じパラメタリゼーションを使用するため他のモデルアンサンブル法に比べてアンサンブルメンバー間の不確定さは比較的小さい。しかし、パラメータ値の与え方には任意性があるため適切なパラメータを見出す必要がある。パラメータの推定法として気候学的値を推定する方法やモデルバイアスを修正する方法 (Annan and Hargreaves 2004; Beak et al. 2006) があるが、パラメタリゼーションにおける最適なパラメータ値を直接推定し、その1か月予報での有効性を調べた研究はない。

本研究の目的は、モデル誤差を持つような不完全な予報モデルにおいて、1か月予報の改善に有効なモデルアンサンブル (マルチパラメータ) 法を考案することである。

2 本研究の手法

予報変数の EnKF と予報変数とパラメータを併せた EnKF から成る (図1)。ここで言うパラメータとはモデルのパラメタリゼーションにおけるパラメータのことである。

まず、予報変数の EnKF のサイクルについて述べる。後述の予報変数とパラメータを併せた EnKF により求められる M 個のパラメータ値のそれぞれに対して、初期値 N メンバーの予報を行う。つまり、この予報は初期値アンサンブル予報とモデルアンサンブル予報を組み合わせた初期値・モデルアンサンブル予報 ($N \times M$ メンバー) となる。この予報値 $x_{ij}^f(t)$ ($i = 1, 2, \dots, I; j = 1, 2, \dots, J$; ここで I は初期値メンバー数、 J はパラメータメンバー数を表す。) を異なるパラメータ間で平均した

$$\hat{x}_i^f(t) = \frac{1}{J} \sum_{j=1}^J x_{ij}^f(t) \quad (i = 1, 2, \dots, I) \quad (1)$$

を EnKF へ取り込み、予報変数の観測値を用いて予報変数の解析値と次の予報のための N メンバーの初期値を作成する。作成された N メンバーの初期値を用いて、前回と同様に初期値・モデルアンサンブル予報を行い再び EnKF を行う。このような解析・予報サイクルを繰り返し最適な予報変数の解析値と摂動を求めていく。

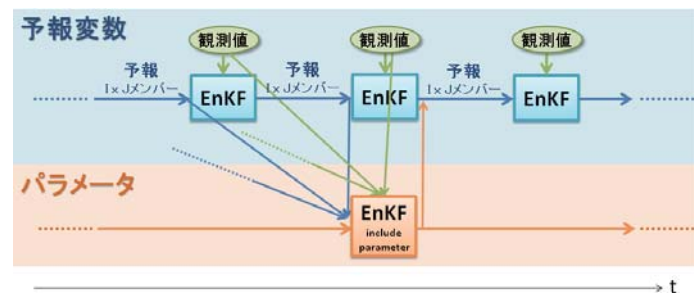


図 1: 本研究の手法の模式図

次に、予報変数とパラメータを併せた EnKF について述べる。これは前述の予報変数の EnKF よりも長い時間サイクルで行う。その目的はサンプリングエラーおよび予報変数の短周期成分を除去することである。前述の予報値 $x_{ij}^f(t)$ を異なる初期値間で平均し、さらに時間平均した

$$\bar{x}_j^f(t) = \frac{1}{T} \sum_{t'=t-a\Delta T}^t \left(\frac{1}{I} \sum_{i=1}^I x_{ij}^f(t') \right) \quad (2)$$

$(j = 1, 2, \dots, J)$

と推定したいパラメータ $p_j(t)$ をそのまま EnKF へ取り込み、時間平均した予報変数の観測値を用いてパラメータの解析値とパラメータの摂動を作成する。ここで ΔT は予報変数の EnKF における同化の時間間隔で、 a は平均する予報値のサイクル数である。パラメータが推定される度に予報モデルのパラメータ値を更新していく。

本手法の最大の特徴は、時間変動する最適なパラメータ値が動的に推定可能であることと同時にモデルアンサンブル予報のためのパラメータ摂動が得られることである。ここで説明した一連の枠組みをまとめて pEnKF と呼ぶことにする。

3 使用するモデル

3.1 完全モデル

以下の Lorenz'96 モデル (Lorenz 1996) をモデル誤差を含まない完全モデルとする。

$$\frac{dx_m}{dt} = x_{m-1}(x_{m+1} - x_{m-2}) - x_m + F - \frac{hc}{b} \sum_{n=1}^N y_{m,n} \quad (3)$$

$$\frac{dy_{m,n}}{dt} = cby_{m,n+1}(y_{m,n-1} - y_{m,n+2}) - cy_{m,n} + \frac{hc}{b} x_m \quad (4)$$

ここで、添字 $m = 1, 2, \dots, M$, $n = 1, 2, \dots, N$ は格子点を表し、 h, c, b はローレンツパラメータ、 F は強制である。 x と y は共に周期境界条件 ($x_{m-M} = x_m$, $x_{m+M} = x_m$, $y_{m+M,n} = y_{m,n}$, $y_{m,n-N} = y_{m-1,n}$) を満たす。

3.2 不完全モデル

モデル誤差を持つような不完全モデルを以下のように作成する。(3) 式の右辺最後の項

$$-\frac{hc}{b} \sum_{n=1}^N y_{m,n} \quad (5)$$

を x_m とパラメータ係数 α, β を用いて表した項 $\alpha + \beta x_m$ (6)

に置き換えることによって得られる

$$\frac{dx'_m}{dt} = x'_{m-1}(x'_{m+1} - x'_{m-2}) - x'_m + F + \alpha + \beta x'_m \quad (7)$$

を不完全モデルの式とする (Smith 2000; Orrell 2003)。ここで、「'」は不完全モデルにおける変数であることを示す。

y は x に比べて格子間隔が狭く、小さな振幅で速く変動することから、(5) を (6) に置き換える操作は y のパラメタリゼーションと考えることができる。これは、大気大循環モデルに当てはめて考えると、 x は直接計算可能な予報変数を表し、 y は直接計算出来ないサブグリッドスケールの現象の変数を表していると思えることができる。

3.3 設定

本研究では、変数の数を $M = 8$, $N = 4$ とし、強制およびローレンツパラメータの値は $F = 10, h = 1$ (Smith 2000; Orrell 2003; Roulston et al. 2003) とした。 c と b は

$$c = b = 2.5 \sin\left(\frac{2\pi t}{T}\right) + 7.5 \quad (8)$$

の形で与えた。これは完全モデルと不完全モデルの間のモデルの不完全さに明瞭な長周期的な変動を加えるためである。今後、特に明記しない限り周期 $T = 90$ 日¹の結果を示す。上記の条件下ではいずれも解はカオス的な振舞いをすることを確認している。また、積分の時間間隔は x, y に対してそれぞれ 0.005, 0.0005 で、4 次の Runge-Kutta 法を用いて積分を行った。

¹誤差の成長速度の観点からこのモデルでの時間 0.2 が現実大気の 1 日に相当することが知られている。今後はこの現実大気の時間に換算した時間を示す。

4 実験設定

完全モデルにおける単一メンバーの時間発展を真値として定義し、その真値に平均 0、標準偏差 $\sigma_x = 1$ 、 $\sigma_y = 0.1$ のガウシアンノイズを加えたものを観測値とした。

使用した EnKF の手法は逐次的アンサンブル平方根フィルタ (Serial EnSRF, Whitaker and Hamill 2002) である。観測は全ての格子点で得られるものとした。予報変数の同化の時間間隔は 0.05 で、これは現実大気における 6 時間を想定している。また、EnKF を安定して動作させるために共分散行列の局所化 (Gaspari and Cohn 1999) を導入した。

予報変数の EnKF で得られる初期値を 1 か月先まで延長することにより 1 か月初期値アンサンブル予報または初期値・モデルアンサンブル予報を行った。

5 結果

最初に初期値アンサンブル予報のみの結果を示した後、pEnKF を導入した結果を示す。

5.1 初期値アンサンブルのみ

図 2(a) は、完全モデル (細線) と一定のパラメータ値を用いた不完全モデル (太線) における予報期間と予報誤差 (RMS 予報誤差; 2 乗平均平方根予報誤差) の関係を示したもので、図 2(b) は不完全モデルの予報誤差から完全モデルのそれを引いたものである。完全モデルと不完全モデルの予報誤差はそれぞれ 21 日、17 日予報で気候値予報 (破線) に漸近することから予報の限界はその付近であることが分かる (図 2b)。また、完全モデルと不完全モデルの予報誤差の差は 6 日予報付近をピークにして予報期間後半まで値を持ち続けることが分かる。

5.2 初期値アンサンブルのみ (統計的パラメータ)

次に不完全モデルにおいて一定のパラメータ値ではなく時間変動する最適なパラメータ

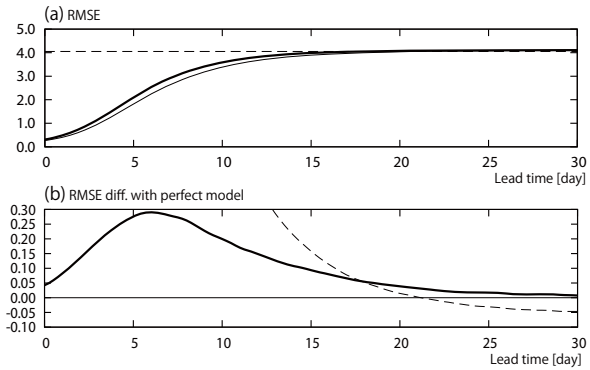


図 2: (a) 完全モデル (細線)、一定のパラメータ値を用いた不完全モデル (太線)、気候値予報 (破線) における予報誤差。横軸は予報期間 (日)。メンバー数は 30 で、30 年平均値。不完全モデルのパラメータ値は (5) 式の線形回帰により求めた。(b) (a) の不完全モデル (太線) および気候値予報 (破線) の予報誤差から完全モデルのそれを引いたもの。

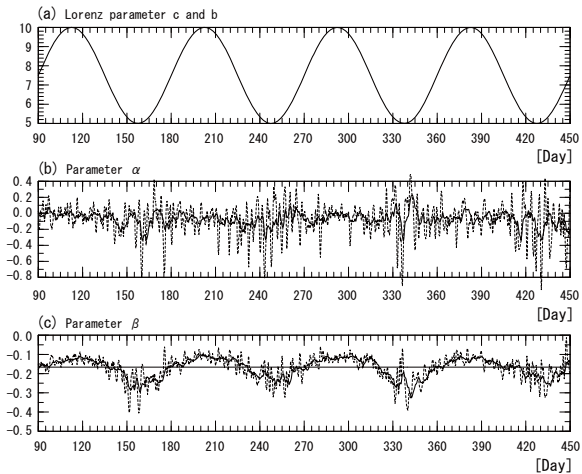


図 3: ある期間における (a) ローレンツパラメータ c, b 、(b) パラメータ α の統計的パラメータ推定値、(c) パラメータ β の統計的パラメータ推定値の時系列。(b) と (c) における太実線、細破線はそれぞれ 5 日、1 日の Time-window で計算された結果。

値を用いることで予報が改善するかどうかを調べる。完全モデルの (5) 式を (6) 式の形で線形回帰することにより統計的に求めたパラメータ値 (統計的パラメータ推定値と呼ぶ) を用いた予報を行う。これは完全モデルが既知の場合に限り計算可能なものであり、現実的には計算不可能な手法であるため、あくまで検証用として用いる。計算する際に用いる時間範囲を Time-window と呼ぶことにする。図 3 はある期間のローレンツパラメータ c, b およびパラメータ

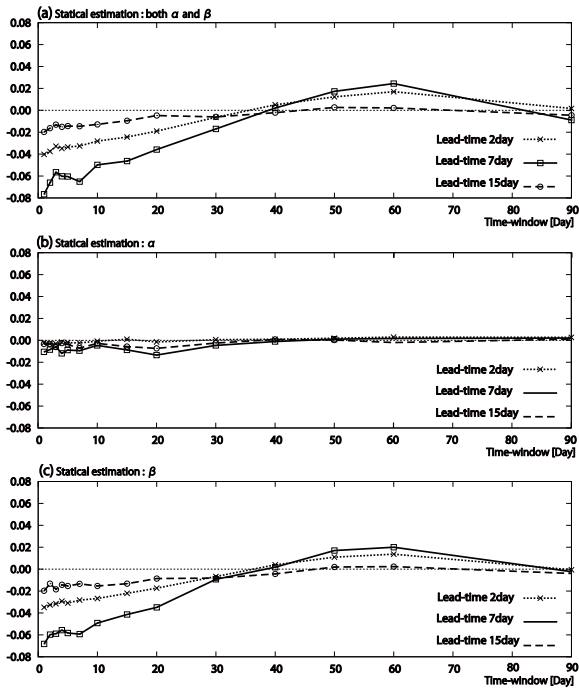


図 4: 統計的パラメータ推定値を用いた予報の予報誤差から一定のパラメータ値を用いた予報のそれを引いたもので、横軸は Time-window(日)。(a) は α と β の両方が統計的パラメータ推定値の場合。(b) は α が統計的パラメータ推定値で、 β が一定値の場合。(c) は β が統計的パラメータ推定値で、 α が一定値の場合。いずれも、2日、7日、15日予報がそれぞれ点線、実線、破線で示されている。

との時系列を示している。はほぼランダム的に変動しているのに対し、は c, b に対応した長周期的な周期性を持つことが分かる。

図 4 は統計的パラメータ推定値を用いた予報の予報誤差から一定のパラメータ値を用いた予報のそれを引いた結果を示している。横軸は Time-window(日)である。図 4(a) より Time-window が短いほど予報が改善し、サンプリングエラーなどの影響を受けていないことが分かる。また、7日予報(実線)で改善の効果が特に大きいことが分かる。のみに統計的パラメータ推定値を用いた場合(図 4b)ではほとんど予報が改善しないのに対し、のみに統計的パラメータ推定値を用いた場合(図 4c)ではと の両方に統計的パラメータ推定値を用いた場合(図 4a) とほぼ同等である。よって予報の改善に対して有効なパラメータは α であることが分かる。これは図 3 で α に比べて β では時間変動に

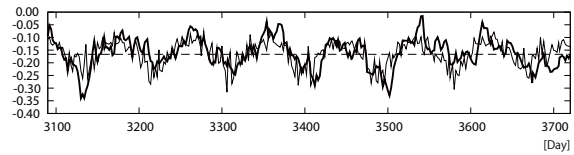


図 5: ある期間における pEnKF により推定されたパラメータ(太線)と統計的パラメータ推定値の時系列(細線)。いずれも α 。pEnKF において、初期値メンバー数 20、パラメータメンバー数 5、総メンバー数 100 である。

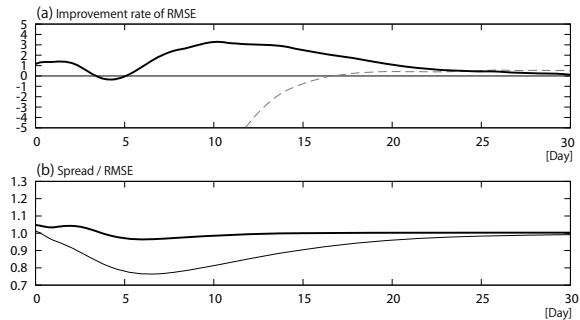


図 6: (a)pEnKF を適用した場合の適用しない場合(100 メンバーの初期値アンサンブル)からの予報誤差の改善率。(b) スプレッドと予報誤差の比。いずれも、太線が pEnKF(初期値メンバー数 20、パラメータメンバー数 5、総メンバー数 100) を適用した場合、細線が pEnKF 非適用の初期値アンサンブル(メンバー数 100) のみの場合、破線が気候値予報の場合。横軸は予報期間(日)。

ランダム的な成分よりも長期的な成分が大きいことに起因すると考えられる。次節では α は一定値とし、 β のみを pEnKF により推定する。

5.3 初期値・モデルアンサンブル(pEnKF)

図 5 は pEnKF により推定されたパラメータのある期間の時系列(太線)を示している。細線は前節の統計的パラメータ推定値の α における時系列であり、これを真に近いパラメータの時系列と考える。両者の長周期的変動がほぼ一致することから、pEnKF によりほぼ 90 日周期の長周期的な変動が推定出来ていることが分かる。また両者の期間を通した相関係数も 0.56 と比較的大きい。

図 6(a) は pEnKF を適用した場合の適用しない場合(初期値アンサンブルのみ)からの予報誤差の改善率を示している。総メンバー数は共

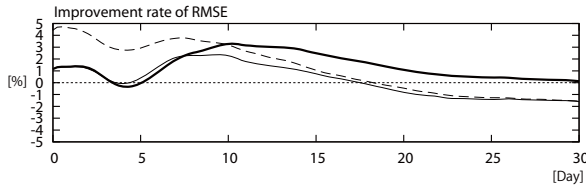


図 7: EnKF を適用した場合の適用しない場合からの予報誤差の改善率。太線が pEnKF(初期値メンバー数 20、パラメータメンバー数 5、総メンバー数 100) を適用した場合で、細線が pEnKF で推定されたパラメータを用いるがモデルアンサンブルは行わない場合(初期値メンバー数 20)、破線が統計的パラメータ推定値を用いた場合(初期値メンバー数 20)。横軸は予報期間(日)。

に同じ 100 メンバーである。5 日予報以降の予報誤差が最大で 3 %程度改善していることが分かる。図 6(b) はスプレッドと予報誤差の比を示しており、この値が 1 に近いほど確率密度分布の推定が適切であると考えられる。pEnKF を適用した場合(太線)ではこの比は適用しない場合(細線)に比べてずっと 1 に漸近していることが分かる。これは主にモデルアンサンブルによる効果である。なお、総メンバー数をさらに増大させていく(図示せず)と、pEnKF を適用しない初期値アンサンブルのみ場合では予報が改善するどころか逆に悪化²するのに対して、初期値・モデルアンサンブル予報の場合では予報は改善する。このことは十分なメンバー数がとれるなら初期値アンサンブルのみよりも初期値・モデルアンサンブル予報の方が有効であることを示している。

pEnKF では最適なパラメータの推定と初期値・モデルアンサンブル予報が併せて行われるため、図 6(a)の結果には最適なパラメータが推定された効果とモデルアンサンブルによる効果の両方が含まれている。そこでこれらのどちらの効果が予報の改善に対して有効であるのかを調べる。図 7 は図 6(a)と同様に pEnKF を適用した場合の適用しない場合(初期値アンサンブルのみ)からの予報誤差の改善率を示している。太線は通常通り pEnKF を適用した場合(図 6a の太線と同一)でパラメータ推定とモデルアン

²モデルの自由度に比べて初期値メンバー数が多すぎためと考えられる。

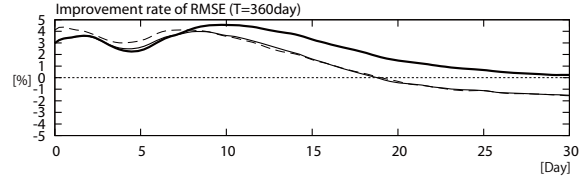


図 8: モデルの不完全さの周期 ((8) 式) が $T = 360$ 日であることを除いては図 7($T = 90$ 日)と同様。

サンプルの両方の効果を含む。細線は pEnKF で推定されたパラメータを用いるがモデルアンサンブルを行わない場合である。初期から 1 週間先くらいまでの予報期間では、モデルアンサンブルあり(太線)となし(細線)の改善率の差がほとんどないことから主にパラメータを推定した効果により予報が改善している。ただし統計的パラメータ推定値を用いた場合(破線)に比べると改善率はかなり小さいことからパラメータの推定精度はそれほど良いとは言えない。1 週間以降の予報期間ではモデルアンサンブルなしの場合(細線)でも予報は改善しているが予報期間の後半に行くにつれてその効果は次第に小さくなる。モデルアンサンブルあり(太線)となし(細線)の改善率の差を見ると、予報期間の後半に行くにつれて大きくなっており、モデルアンサンブルは予報期間の後半ほど有効であることを示している。

最後に、モデルの不完全さの周期をより長くした場合の結果を示す。図 8 は (8) 式において $T = 360$ 日であることを除いては図 7($T = 90$ 日)と同様の図である。モデルアンサンブルなし(細線)の改善率が統計的パラメータ推定値を用いた場合(破線)と同程度であり、 $T = 90$ 日の場合に比べてパラメータの推定精度がかなり良いことが分かる。一方でモデルアンサンブルあり(太線)となし(細線)の改善率の差を見ると、 $T = 90$ 日の場合と同様の傾向を示しており、モデルアンサンブルは予報期間の後半ほど有効である。よって、パラメータの推定は特にモデルの不完全さの周期が長い場合に予報の改善に対して有効であり、モデルアンサンブルはモデルの不完全さの周期にあまり依存せず 1 週間以降の予報の改善に有効であることが分かる。

6 まとめ

EnKF を応用したパラメータを含めた同化・予報システムを構築した (pEnKF)。この最大の特徴は最適なパラメータ推定と初期値・モデルアンサンブル予報を行うことである。モデル誤差をもつ Lorenz'96 モデルにおいて pEnKF を適用した結果、最適なパラメータの推定が可能であり、1 か月予報においても 5 日予報以降で初期値アンサンブルのみ場合よりも予報が改善することが分かった。特に、予報期間の前半ではパラメータ推定の効果、後半ではモデルアンサンブルの効果により予報が改善していると考えられる。

今後は、大気大循環モデルへ本手法を適用しその有効性を調べ行きたい。その中で推定しやすいパラメータとしづらいパラメータの違いをパラメタリゼーションの構造に着目して理解することが必要である。また、パラメータ摂動の構造の解析やより効果的なパラメータ推定法 (例えば、より長期の予報値を用いてパラメータを推定するなど) への改良も行っていきたい。

7 謝辞

今回の研究集会の参加にあたり、東京大学気候システム研究センターより旅費の援助を受けました。感謝致します。図の作成には DCL(地球流体電脳ライブラリ) を使わせて頂きました。

参考文献

Annan, J. D. and Hargreaves, J. C., 2004: Efficient parameter estimation for a highly chaotic system. *Tellus*, **56A**, 520-526.

Baek, S.-J., Hunt, B. R., Kalnay, E., Ott, E., Szunyogh, I. 2006: Local ensemble Kalman filtering in the presence of model bias. *Tellus*, **58A**, 293-306.

Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.*, **99** (C5), 10143-10162.

Gaspari, G. and S. E. Cohn, 1999: Construction of correlation functions in two and three dimensions. *Quart. J. Roy. Meteor. Soc.*, **125**, 723-757.

Lorenz, E. 1996: Predictability - A problem partly solved. *Proc. Seminar on Predictability*, Shinfield Park, Reading, United Kingdom, European Centre for Medium-Range Weather Forecasting, 1-18.

Miyoshi, T., 2005: Ensemble Kalman filter experiments with a primitive-equation global model. Doctoral dissertation, University of Maryland, College Park, 197pp.

Orrell, D., 2003: Model error and predictability over different timescales in the Lorenz'96 systems. *J. Atmos. Sci.*, **60**, 2219-2228.

Roulston M. S., and L. A. Smith, 2003: Combining dynamical and statistical ensembles. *Tellus*, **55A**, 16-30.

Smith, L. A., 2000: Disentangling uncertainty and error: On the predictability of nonlinear systems. *Nonlinear Dynamics and Statistics*, A. Mees, Ed., Birkhauser, 31-64.

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