

非ガウス分布の定量的評価による非ガウスデータ同化

*近藤圭一（気象研究所）、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よく使用されているアンサンブルカルマンフィルタ (EnKF; Ensemble Kalman Filter, Evensen 1994) では、二次の統計モーメントまでを扱った最小分散推定を行う。この結果は、誤差分布がガウス分布であるときに最尤推定と一致する。Kondo and Miyoshi (2019, NPG) はアンサンブル数 10240 からなるアンサンブルデータ同化実験の結果を用いることで、大気における背景誤差の確率分布を詳しく検証した。その結果、図1のように背景誤差は多くの格子点でガウス分布であるものの、熱帯での活発な対流やストームトラック付近での傾圧不安定等に起因する強非線形性が非ガウス分布を発生させることを確認した。またそのような非ガウス性が強い領域では、EnKF の解析誤差が大きくなる傾向があることを示した。このように非線形性が卓越するような現象では、誤差分布が非ガウスとなり解析精度を悪化させ、予測が困難になるという悪循環を生むことにつながる。

このような非ガウス分布に対応するには、粒子フィルタ (PF) が有効である。PF は計算負荷の小さい低次元系では多用されているが、大気モデルのような多次元複雑系では PF に必要な膨大なアンサンブル数 (粒子数) を用意することは、計

算資源の制約から困難とされてきた。しかし近年、Penny and Miyoshi (2016), Poterjoy (2016), Potthast et al. (2018), Kotsuki and Miyoshi (2020) のように限られた粒子数で動作可能となる PF の研究が進んでいる。

本研究では、従来のアンサンブルデータ同化手法を拡張した非ガウスデータ同化手法を提案し、数値天気予報におけるデータ同化の精度向上を図ることを目的とする。

2. 手法

本研究では、大気モデルが持つ強非線形性に起因する非ガウス分布を適切に扱える非ガウスデータ同化手法を開発する。図1で示したとおり SPEEDY では多くの領域で背景誤差はガウス分布である。現実的な NICAM を用いた場合では非ガウス性が強い領域は増えるが、それでも非ガウス性の低い領域は多く存在する。非ガウス性の低い領域では従来通りガウス分布を仮定した EnKF が有効であり、非ガウス性の強い領域に限定して非ガウスデータ同化を導入する方法が考えられる。本研究は、この方針で以下の手順を考案した。

① 観測毎に予報アンサンブルの分布から非ガウス性を KL divergence (KLD; Kullback and Leibler 1951) で評価する。KLD はガウス分布であれば 0、非ガウス分布強度が高いほど大きくなる。

② EnKF で観測を同化する。この際使われる観測誤差分散 $\sigma_{obs, EnKF}^2$ には以下を用いる。

$$\sigma_{obs, EnKF}^2 = \sigma_{obs}^2 / (1 - \alpha)$$

σ_{obs}^2 はもともとの観測誤差分散を表す。 α は KLD に応じて [0,1] 区間から決まり、ガウス分布に近いほど 0 に近づき、非ガウス性が強くなるほど 1 に近づく (図2)。観測誤差が大きくなるため、観測は軽く扱われる。

③ 上記 EnKF の解析アンサンブルを第一推定値として、PF で同じ観測をもう一度同化する。この際使われる観測誤差分散 $\sigma_{obs, PF}^2$ は以下を用いる。

$$\sigma_{obs, PF}^2 = \sigma_{obs}^2 / \alpha$$

②と同様、観測誤差が大きくなり、観測は軽く扱われる。

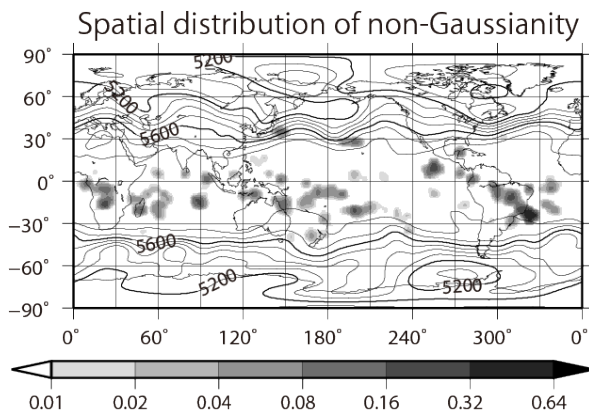


図1 ある年の2月22日06 UTCにおけるSPEEDYのモデル4層目 (~500 hPa) の気温について、誤差分布の非ガウス性強度を示す。数値が大きいほど非ガウス性が強い。等値線は500 hPa 高度。Kondo and Miyoshi (2019)の Fig. 4 より一部改変。

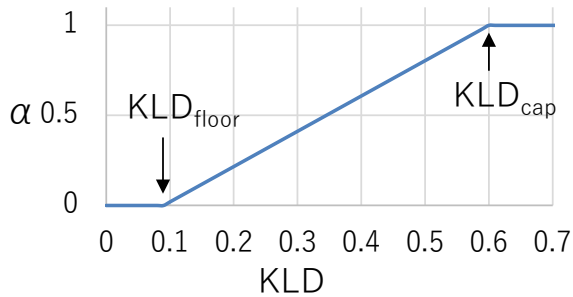


図2 KLD と α の関係。

例えば、ある観測について $KLD=0$ であれば、完全にガウス分布に従う。このとき $\alpha=0$ となり、PF では同化されない。非ガウス性が強くなると KLD が大きくなり、 α も大きくなる。 α が 1 に近づくほど EnKF の効果が小さく、また PF の効果が大きくなる。EnKF に PF を組み合わせているためハイブリッド同化手法と呼ぶ。

手順①②のみを実行する場合は、EnKF システムを変えることなく、各観測の観測誤差を非ガウス分布強度に応じて動的に膨張させるのみでよい。本研究ではこの手法を動的観測誤差膨張手法と呼び、簡易的な非ガウスデータ同化手法の 1 つとする。

少ないアンサンブル数ではサンプリング誤差により KLD の評価は困難である (Kondo and Miyoshi 2019)。そのため、図 2 の KLD_{floor} は、サンプリング誤差を推定するため 100 万個のガウス乱数から統計処理を行うことで求める。 KLD_{cap} は SPEEDY を用いた 10240 メンバーデータ同化実験の結果から経験的に求める。これにより、少ないアンサンブル数で KLD による非ガウス性強度の判定が可能となった。

3. 実験設定

本研究では、簡易全球大気モデル SPEEDY (T30/L7, Molteni 2003) に局所アンサンブル変換カルマンフィルタ (LETKF; Hunt et al. 2007) を適用した SPEEDY-LETKF システム (Miyoshi 2005) をベースとして使用する。アンサンブル数は 80 とする。観測はラジオゾンデ観測 (線形観測演算子 $H(x) = x$ 、観測変数は U, V, T, Q, Ps) とモデル格子点の 2×2 に 1 点の一樣観測 (非線形観測演算子 $H(x) = \log(|x|)$ 、観測変数は T, Q) を 6 時間毎に同化する観測システムシミュレーション実験 OSSE を行った。実験期間は 1 年 2 ヶ月

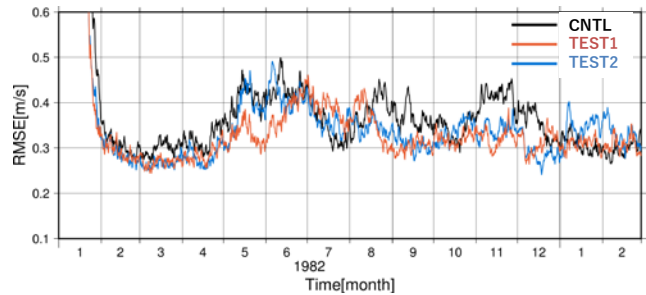


図3 モデル4層目 (~ 925 hPa) における東西風の解析誤差 RMSE 時系列。黒線は CNTL、赤線は TEST1、青線は TEST2 を示す。

である。CNTL 実験では通常の LETKF を、TEST 1 実験ではハイブリッド同化手法を、TEST2 実験では動的観測誤差膨張手法をそれぞれ使用している。

4. 結果と考察

図 3 は、モデル 4 層目 (~ 500 hPa) における東西風の解析 RMSE の時系列である。TEST 1, TEST2 とともに安定して動作し、CNTL 実験より概ね改善している。RMSE の対 CNTL 改善率は TEST1 で 8.9%、TEST2 で 6.2% である。

ラジオゾンデ観測のみを同化した別の実験を行ったところ、CNTL、TEST1、TEST2 で解析精度に大きな違いは確認されなかった (図略)。このことから、図 3 の結果は、非線形観測演算子を用いると観測空間における非ガウス強度が大きくなり、非ガウスデータ同化の効果が現われるためだと考えられる。

これらの結果は、他の変数及び高度でも同様の結果であった。

5. まとめ

本研究では、予報アンサンブルの非ガウス性の強度に応じて PF を限定的に導入することで、EnKF とハイブリッド型の新しい非ガウスデータ同化システムを考案した。SPEEDY モデルを使った理想実験を行い、本手法が少ないアンサンブル数でも安定して動作し、非線形観測演算子を用いた場合に従来手法よりも解析誤差が小さくなることを示した。特に動的観測誤差膨張手法は、観測を非ガウス性強度に応じて膨張させる簡便な手法であるが、解析精度向上インパクトは大きい。

今後はより現実的な数値天気予報システムを用いて検証していきたい。