令和2年度　卒業論文

# Lorenz-96モデルを使った粒子フィルタのデータ同化インパクトの研究


筑波大学生命環境学群地球学類

地球環境学主専攻

**201913525**

## 赤見彰一


**2021年1月**

# A Study on the data Assimilation Impact with Particle Filter on the Lorenz-96 Model


**201913525**

**Shoichi AKAMI**

College of Geoscience(Geoenvironmental Sciences)

School of Life and Environmental Sciences

University of Tsukuba


in Partial Fulfillment of the Requirements

for the Degree of Bachelor of Science

**January 2021**

# 目 次

# Lorenz-96 モデルに対する 粒子フィルタのデータ同化インパクトの研究

赤見　彰一

## 要旨

粒子フィルタ (PF) を大気モデルで機能させる為の条件を明らかにする事は, 近年のデータ同化という研究分野において, 最も重要な問いの 1 つである. データ同化 (DA) という手法が 1950 年代に実現されてから現在に至るまでの間に, アンサンブルカルマンフィルタ (EnKF) や 4 次元変分法 (4D-Var) といった様々な手法が考案されてきた. しかし, 既存のデータ同化手法は全て予報誤差がガウス分布になるという仮定を置いており, 大気の非線形性が強い場合はこの仮定が満たされなくなるため, 解析精度が低下する.PF はこの様な仮定をしないため, 非線形性が強い顕著現象に対しても有効だが, 大量のアンサンブル数を必要とする。したがって, 長らく現業大気モデルの様な多次元系では機能しないとされてきたが, 近年になって機能できる可能性が浮上してきた. 本研究では,Lorenz-96 モデルに既存手法として EnKF の一種の逐次平方根フィルタ (Serial EnSRF)・観測摂動 (PO) 法, 及び粒子フィルタ (PF) を実装し, チューニングパラメータを変化させて数値実験を繰り返す事で,PF が大気モデルで機能し, かつ既存手法よりも解析精度が上回る条件を明らかにする.

結果として PF が機能したとしても、EnKF より解析精度が低くなった.PF は, 粒子数 (アンサンブル数) が少ないと機能せず、かつアンサンブルの退化が生じうる。PF は局所化スケールや膨張係数に対して非常に敏感であり, チューニングパラメータを適切な値にしないとフィルタ発散に陥る。したがって、現業予報で用いる為には、それらの動的推定を実現する事が必要になる。しかし、PF は非ガウス性によって解析精度が低下し難いため、海洋の様に観測点が少ない領域では EnKF より有効になる可能性がある。

キーワード：データ同化, 粒子フィルタ,Lorenz-96 モデル, 逐次平方根フィルタ, 摂動観測法, 共分散膨張, 局所化, カルバックライブラー情報量

A Study on the data Assimilation Impact

with Particle Filter on the Lorenz-96 Model

Shoichi AKAMI

## Abstract

Clarifying the conditions under which Particle Filters(PF) can works in atmospheric models is one of the most inportant questions in recent research the field of Data Assimilation(DA) . During the period from the beginning of data assimilation in the 1950s to the present, various methods such as the Ensemble Kalman Filter(EnKF) and the 4-dimensional Variational method(4D-Var) have been devised. However, all of the previously data assimilation methods use the assumption that the forecast error is Gaussian. Therefore, the case of strong nonlinearity in the atmosphere, this assumption is not satisfied, and the analysis accuracy is degraded. PF doesn't use Gaussian assumptions and is therefore effective for pronounced phenomena with strong nonlinearity. However, because the PF requires a large number of ensemble members, it has long been considered that PF is unworkable in high dimensional systems such as operational atmospheric models. But recently the possibility of that it can work has emerged. In this study, we implement the EnKFs such as the Square Root Filter(SRF) and the Perturbed Observation(PO) method, and the PF with localization using the Lorenz-96 model, and repeat numerical experiments with different tuning parameters to clarify the conditions under which the PF can works in an atmospheric model and outperform the EnKFs in terms of analytical accuracy.

As a result, even if PF worked, the analysis accuracy was lower than EnKF.The PF doesn't work in the case of number of effective particles (ensemble number) is small, and ensemble degeneration can occur. PF is very sensitive to localization scale, inflation coefficient, and if not given the proper tuning parameters, filter divergence is generated. Therefore, it is necessary to realize the adaptive estimation of these parameters in order to use PF in

operational forecasts. However, PF has the potential to be more effective than EnKF in regions with few observation points such as the ocean, Because it is difficult to degrade analysis accuracy by non-Gaussian in PF.

Keywords: Data Assimilation, Particle Filter, Lorenz-96 model, Serial Square Root Filter, Perturbed Observation method, Covariance inflation, Localization, Kullback–Leibler divergence

# 図目次

# 第1章　はじめに

## 1.1　大気のカオス性とデータ同化

本項では、これから本研究の内容を理解していく上で必須となる、大気のカオス性やデータ同化の概念について、近藤 (2009) に基づき簡単に述べていく。

まず初めに、大気の振る舞いにはカオス性というものが存在する。これは大気の状態変化が非線形微分方程式によって記述される事に起因しており、非線形微分方程式は解析的に解く事が不可能なため、差分法を用いて数値的に解くしかない。しかし、差分法によって数値解を近似的に得ようとすると、そこには必ず誤差が含まれてしまう。この誤差がもたらす影響は時間経過と共に増大するため、最初の段階では僅かな誤差しかなかった複数の初期値から得られる解が、最終的に全く異なるものになってしまう事がある。これをバタフライ効果と言い、蝶の羽ばたきが起こす僅かな気流の乱れが大嵐になりうる、と比喩的に語られる事がある。

したがって、たとえ数値予報モデルが完璧であったとしても、バタフライ効果によって初期値の僅かな誤差が時間と共に発達し、最終的に数値解が大幅に異なったものになってしまうため, 長期間にわたる気象予測を正確に行う事は非常に困難である. この様な不確かさを考慮し、より高精度の気象予測を可能にする手法の 1 つにアンサンブル予報が挙げられる. これは、初期値が僅かに異なる複数の数値予報を行い、その結果を統計的に処理する事で, 大気のカオス性を考慮した確率的な予報を可能にするものである. また、この手法では複数の予報を行う事でアンサンブルメンバーの広がりから予報の信頼性を予め見積もる事もできる. しかし、複数の予報を行う事で生じる計算コストは非常に重く, 基本的には数値予報モデル

の自由度よりも大幅に少ない数のアンサンブルメンバーしか作成する事ができない. たとえば、大気大循環モデルの自由度はおよそ $N = O(10^7〜10^8)$ に達するが, それに対してアンサンブルメンバーはおよそ数百〜数千しか作成できない. 更に、アンサンブル予報は計算機資源を大量に消費する都合上, 決定論的予報を行うモデルと同等の解像度を維持する事は難しく, より細かいスケールの現象を再現する事はできない.

気象予測の精度を向上させるもう 1 つの手法としてデータ同化が挙げられる. これは、予報値と観測値の重み付け平均から得られる、より高精度の初期値を数値予報モデルに与える事で、大気のカオス性によって時間と共に増大する誤差の発達を抑え, より高精度の予測を可能にする統計的数理に基づいた手法である. 数値予報モデルの予報値にはパラメタリゼーションなどに起因する予報誤差が含まれ, 観測値には測器の不確かさなどによる観測誤差が常に含まれる. データ同化では予報誤差と観測誤差の分散共分散を用いて観測値と予報値の重み付け平均を計算し、そうして得られる観測の情報を反映したより尤もらしい解析値を次の時刻の初期値として数値予報モデルに与える事で、大気のカオス性に起因する予測精度の低下を抑える事を可能とする. 簡単な例を挙げると、気温の予報値が 20 ℃であったのに対し、観測値は 30 ℃であった。本来は決して知る事ができないが、この時の真値を 25 ℃とする。すると、予報値と観測値はそれぞれ± 5 ℃の予報誤差と観測誤差を含んでいる事になる。この場合、予報値と観測値の重みは 1:1 と等しいため、単純に相加平均すると 25 ℃となり、これが解析値に相当する。すると、解析値と真値は等しくなり、予報に観測の情報が反映された事で誤差が無くなり、これを次の時刻の初期値として数値予報モデルに与えれば、大気の非線形性に起因する初期値鋭敏性の影響を抑えた、より確実な予測を行う事ができる。この様にデータ同化とは言い換えれば、観測とモデルを融合し, モデルが現実から離れて独り歩きしない様に留めておく為の手法という事になる.

## 1.2 データ同化の研究史

本項では, 本研究を行うにあたりどの様な経緯を経てデータ同化という技術が発達し, かつその中で何故粒子フィルタに関する研究が行われる様になったのか紹介する為に, 露木 (2008) に基づきデータ同化の研究史について述べていく.

大気の状態変化を数値的に計算して天気を予報するという考え方は1922年にL. F. Richardsonによって提唱された. これは大気の状態変化とはすなわち物理現象であり, ある時刻における観測値を用いてその物理現象を記述する微分方程式を数値的に解く事ができれば, 将来の天気を決定論的に予報する事ができるという考え方であり,「Richardsonの夢」と呼ばれる.Richardsonはこの考えに基づいて手計算による初めての天気予報を試みたものの失敗に終わった. これは当時の計算技術には制限があり僅か6時間の予報に2カ月も要した事に加え, 数値解析の手法に様々な問題を含んでいた為である.1950年代に入りJ. L. Neumannが中心となってコンピュータによる数値予報が研究される様になるとこの問題は解決され,1954年には遂にC. A. Rossbyの協力によってスウェーデン王国空軍気象サービスが現業気象予報の実施に成功した.

ある現象の時間変化を数学的に記述した系を力学系と呼び, 現在の気象予報は気象現象を再現した数理モデルの計算をスーパーコンピュータを用いて行う事で実施されている. より精緻な予報を実現するには気象研究の進展や気象現象を数値予報モデルで正確に再現する為の手法を新たに開発する必要がある. また数値予報モデルの空間解像度や時間解像度を向上させれば地形や大気の状態変化をより精密に再現できる様になるため, 気象予報の精度は向上する. しかし空間解像度を2倍にするとCFL条件により計算コストはおよそ10倍のオーダーで増加するため, 気象予報の精度向上には計算技術の発達も必要不可欠である. 現在の数値予報モデルは離散化の方法に基づいて格子点モデルとスペクトルモデルの2種類に大きく分類する事ができる. 前者は高解像度化に伴う計算コストの増加を$O(N^2)$のオーダーに抑える事ができるが等経度を格子点で区切っているため, 極域と熱帯の格子間隔に差が生じて取り扱いが困難になってしまう. 対して後者は波の重ね合わせによって大気の状態を再現しているため力学過程の計算をより高精度に行う事ができるが, 物理空間と波数空間の変換に時間を要する上に高解像度化に伴う計算コストの増加が$O(N^3)$のオーダーにも達するという欠点がある. この様に数値予報モデルの解像度には限界があり, 大気の振る舞いを完全に再現する事はできない. また, 気象予報に欠かせない観測データは空間的に偏在しており, 特に海洋や南半球で不足しているため、それらの領域では特に予報を始める為に必要となる初期値に不確かさが残る. 上述した様に、初期値の誤差はたとえそれが僅かであっても大気のカオス性によって増大し、気象予測の精度を低下させる。こうした数値予報モデルの欠点を補い、

気象予測の精度を向上させる手法の 1 つがデータ同化である.

データ同化という手法が 1950 年代に考案されてから現在に至るまで, その手法は観測値をどの様にして数値予報モデルに与え, より高精度の初期値を作成するかという観点から発展を遂げてきた. 当初は実況天気図の値が数値予報モデルのどの格子点上に相当するか人間が判断する事で, 観測値を数値予報モデルに反映させていた. しかし、この方法では時間を要してしまうため, 空間的に不規則に分布した観測値から数値予報モデルの規則的な格子点上の値をコンピュータで算出する方法が考案された. これは客観解析と呼ばれ, 人間が直接行う主観解析とは区別された.

最初の客観解析は関数当てはめ法 (Panofsky 1949) であり, これは実際の観測値の分布に上手く適合する様な関数を探し, それを用いて数値予報モデルの格子点上に高精度の初期値を空間内挿するものであった. この手法は気象庁において上部成層圏解析に 2001 年まで使用されていた (村上 1997) が, 観測によって予報を修正する訳ではないため厳密にはデータ同化とは呼ばない. これに対して観測値に適当な重みを与えて予報値を修正する手法を逐次修正法 (Eliassen 1954) と呼び, 更にこの重みの求め方を経験的なものに頼るのではなく, 統計的に推定する様に発展させたものが最適内挿 (OI: Optimum interpolation; Gandin 1963) 法である.1960 年代に入ると数値予報モデルの予報値を第一推定値としてデータ同化を行う事が提案され (Gilchrist and Gressman 1954) , そうする事で解析予報サイクルを繰り返す内に観測データの密な領域の情報が次第に観測データの疎らな領域へ空間的かつ時間的に伝播し, 解析精度を向上させる事が発見された (Thompson 1961) . この様にデータ同化によって修正された情報が将来にも伝わるタイプの手法を 4 次元データ同化と呼ぶ. 更に 1970 年代になると観測システムが大きく発展し, 全球的な気象観測網が整備された. 特に極軌道衛星は観測地点を変えながら時間的に連続したデータを取得するため, 非定時の観測データを同化する手法として直接代入法 (Charney et al. 1969) やナッジング (Hoke and Anthes 1976) が考案された.

1980 年代に入るとデータ同化の手法は更に大きく発展し,3 次元変分法 (3D-Var: 3-dimensional Variational method; Parrish and Derber 1992) やカルマンフィルタ (KF: Kalman fitlter; Kalman 1960) が新しく研究される様になった.OI 法では数値予報モデルの変数と線形関係の

観測データしか同化できなかったのに対し,3D-Varでは数値予報モデルでシミュレートできる物理量ならば原理的にどの様な観測データでも同化する事ができるため, たとえば衛星輝度温度データを同化して台風の予測精度を向上させる事も可能である (Honda et al. 2019) .1990年代には3D-Varを時間軸方向にも拡張した4次元変分法 (4D-Var: 4-dimensional variational method; Liu and Zou 2001) が開発された. 数値予報モデルは地形強制や熱源といった外力や, 大気のカオス性すなわちモデル誤差の影響を受ける事で予測精度が低下する. そこで4D-Varでは評価関数を用いて数値予報モデルから得た解析値が第一推定値や観測値からどれだけ離れているか評価した上で, アジョイントモデルによって時間を遡り評価関数が適切な値になる様に解析値に修正を加えるというプロセスを繰り返し, 最もバランスの取れた最適な解析値を求める. こうする事で観測値によって予報値を修正する際に大気の運動に関わる諸法則をそのまま取り込む事でより高精度の解析値が得られるという仕組みになっている. しかし4D-Varは数値予報モデルを繰り返し実行するため計算コストが膨大であり, かつ数値予報モデルが改善される度にアジョイントモデルを作り直さなければならないため, 非常に維持・開発コストがかかるデータ同化手法である.

KFは予報誤差と観測誤差がガウス分布で, 誤差の時間発展が線形である時に最適な解析値を求めるが, 数値予報モデルの様な非線形システムを扱う為には何らかの工夫を加える必要がある. 拡張カルマンフィルタ (EKF: Extended Kalman Filter; Jazwinski 1970) は全ての状態変数を線形化する事でKFを非線形システムに適用している. しかし数値予報モデルの自由度は膨大なため,EKFをそのまま適用しようとすると逆行列計算が難しく解析値の推定が不安定になるという問題が生じる. そこでアンサンブルカルマンフィルタ (EnKF: Ensemble Kalman Filter; Evensen 1994) ではEKFの予報誤差共分散行列をアンサンブル予報を用いて近似する事でこの問題を回避している.EnKFはアンサンブル予報から外力やモデル誤差を考慮した予報誤差, すなわち流れに依存した予報誤差共分散行列を推定している. この方法ならば4D-Varとは異なりアジョイントモデルは必要なく, 維持・開発コストは小さい. またEnKFはアンサンブル予報から得た予報誤差共分散行列を用いてデータ同化を行い解析値を求めるが, その際にアンサンブルアップデートによって得られる解析誤差をアンサンブル予報の初期摂動とする. こうする事で解析値を求める度に初期摂動を与える必要が無く, アンサンブル予報とデータ同化を同じシステムで実行できるため, 現業数値予報に向いた非常に運用効率の高い解析予報サイクルを実現できる.

EnKF は次の時刻のアンサンブル予報を行うために必要となる初期摂動を作成する過程, すなわちアンサンブルアップデートの方法によって大きく 2 つに分けられる.1 つは予報だけではなく観測にも摂動を加えてアンサンブルアップデートを行う事から摂動観測 (PO-EnKF: Perturbed Observation method; Houtekamer and Mitchell 1998) 法と呼ばれる. もう 1 つは EKF の解析誤差共分散行列の平方根を直接計算してアンサンブルアップデートを行う事から平方根フィルタ (SRF: Square Root Filter) と呼ばれる.PO-EnKF ではそれぞれのアンサンブルメンバーに対して独立した解析予報サイクルを実施する. この時に同じ観測データを用いると観測誤差が考慮されない為に解析誤差が過剰に小さくなってしまう. そこで整合性を保つ為に観測に摂動を加えているが, これは新たなサンプリングエラーを導入させてしまうため, アンサンブル数が少ない場合は観測に摂動を与えない SRF の方が高性能になりやすい傾向がある.SRF では解析誤差共分散行列が解析誤差行列とその転置行列の積によって表される事を利用してアンサンブル予報の初期摂動を求めている. アンサンブルメンバーとなる解析誤差にはそれぞれ観測が同化されており, あるメンバーの解析誤差は各メンバーの予報誤差を少しずつ足し合わせたものになっている. これを概念的に説明すると予報誤差の広がりを変換行列によって解析誤差の広がりにしているという事になる.SRF の中でも最もシンプルな逐次平方根フィルタ (Serial EnSRF: Serial Ensemble Square Root Filter; Whitaker and Hamill 2002) はアンサンブルメンバーが少ない場合でも比較的高精度の解析値を得られるが, 逐次的に観測を 1 つずつ同化するため, 並列化が難しく計算効率が低いという欠点がある. 以上の様に,EnKF はアンサンブルアップデートをはじめとした手法の違いによって様々なものに分類される. その中でも, 気象庁の現業数値予報にも利用されている局所アンサンブル変換カルマンフィルタ (LETKF: Local Ensemble Transform Kalman Filter; Hunt 2005) は SRF の一種だが, 局所化を EnKF のアルゴリズムに取り入れる事で SRF の欠点である並列計算効率を非常に高める事に成功している.

KF はモデル誤差が無い場合でもモデルの非線形性に起因して誤差共分散を過小評価する. すると観測の重みが過小評価され, 予報が観測から離れて独り歩きを始めてしまうフィルタ発散と呼ばれる現象が発生する. これを防ぐ為には誤差共分散を人為的に大きくする必要があり, この技術を共分散膨張 (Pham et al 1998) と呼ぶ. これは EnKF における重要なチューニングパラメータの 1 つであり,Miyoshi (2011) は KF を用いる事で共分散膨張の動的推定を

可能にした. しかしこの手法には欠点があり, 観測誤差が明らかな大気モデルでは機能するが, そうでない現実大気の場合は機能しない. そこで Li et al. (2009) は観測誤差も KF を用いて動的推定し, 解析予報サイクルを繰り返す事で現実大気においても共分散膨張の動的推定が可能である事を示した.

データ同化の解析精度はモデルの性質やチューニングパラメータによっても変化するため,4D-Var と EnKF のどちらがより優れているかという事は一概には言えない. 気象庁では長らく 4D-Var が現業数値予報モデルに利用されてきたが,2020 年には 4D-Var と EnKF のハイブリッド方式が採用される様になった. 誤差の時間発展が線形でモデル誤差が無ければ, 充分長い同化ウィンドウを取った 4D-Var と KF は同値になる (露木 1997) . 対して, 無限大のアンサンブルメンバーを作成する事ができれば EnKF と KF は同値になる. しかし実際には大気モデルは非線形であり, 計算機資源にも限りがあるため完璧な精度でデータ同化を行う事はできない.

ただし,4D-Var も EnKF も予報誤差がガウス分布になるという仮定を置いているため, その仮定が満たされない集中豪雨などの顕著現象においては, データ同化の解析精度が低下する.Kondo and Miyoshi (2019) では,10240 メンバーを用いた数値実験を行う事で, 顕著現象の発生に繋がる大気の擾乱が予報誤差分布の非ガウス性を生み出し, それによってデータ同化の解析精度が低下する事を明らかにした. この様な状況下で, 新たに粒子フィルタ (PF: Particle Filter; Kitagawa 1996) というデータ同化手法が考案された.PF は既存のデータ同化手法と異なり, ガウス分布の仮定を置かないため, 予報誤差分布が非ガウス分布になる顕著現象に対しても有効だと考えられているもっとも,.PF は十分な粒子数がないと機能しないため, 大気モデルの様な多次元系には実装できないと考えられてきた. しかし, 近年になって条件次第では PF を大気モデルに実装できる可能性が浮上してきており, 最近ではその条件を明らかにする為の研究が盛んに行われている.

論文の構成としては第 2 章で Lorenz-96 モデルの解説に加え, 共分散膨張や局所化の説明と KF・EnKF・PF の導出を行った. 第 3 章では Lorenz-96 モデルを用いた双子実験によって PO-EnKF・Serial EnSRF・PF の解析精度を比較し, その結果について考察した. 第 4 章では研究全体の考察とそこから得た結論を記した.

## 1.3 目的

これまでデータ同化という手法が1950年代に実現されてから現在に至るまでの研究史を述べてきたが, 既存の手法は全て予報誤差がガウス分布に従うという前提の上に成り立っている. そのため熱帯の積雲対流や低気圧の周辺といった擾乱が激しい状況では, 大気の非線形性が強まる事で予報誤差が非ガウス分布になり, データ同化の精度が低下するという問題が生じる. 近年は平成30年7月豪雨や令和元年東日本台風といった災害が増加傾向にあり, 被害を軽減する為には集中豪雨をはじめとした顕著現象の予測精度を向上させる事が必要不可欠である. そこで本研究では粒子フィルタ (Particle Filter: PF) というデータ同化手法を用いる事で数値予報の精度がどれほど向上されるか明らかにする.

PFはKFの一種であり予報誤差がガウス分布に従うという前提を置かないため, 上述した様な顕著現象に対しても有効な可能性が示唆されている. 解析手法としてはLorenz-96モデルという簡易モデルにPFと既存のデータ同化手法を実装し, 数値実験を行う事で両者を比較する. また将来的に数値予報モデルにPFを実装する事を念頭に置き, 大気モデルにおいてPFを正常に機能させ, かつ予測精度を向上させる為に必要となる条件を調べる. 更にそれぞれのデータ同化手法の特徴についても考察する.

# 第2章　解析手法

## 2.1　Lorenz-96 モデル

Lorenz-96 モデル (Lorenz 1996 Model; Lorenz 1996) は一定の緯度圏上に存在する任意の物理量の時間発展を表す簡易気象モデルであり, その方程式系は

$$\frac{dx_i}{dt} = (x_{i+1} - x_{i-2})x_{i-1} - x_i + F, \quad (i = 1, ..., N) \tag{2.1}$$

となっている. 添字 i = i,...,N は自由度を表し, 周期境界条件を

$$x_i = x_{N+i}, \quad (-\infty < i < \infty) \tag{2.2}$$

とする事で cyclic model にしている. 右辺第 1 項は移流項, 第 2 項は拡散項, 第 3 項は外力項である.

Lorenz-96 モデルは自由度すなわち格子点数 N と外力項 F を任意の値に設定する事で, モデルの非線形性すなわち誤差の発達率を調節する事が可能である. また比較的単純な方程式系によって表されることから計算コストが少なく, モデルの自由度を超える大量のアンサンブルメンバーを作成する事ができる. こういった特徴からアンサンブルメンバーが多い場合に効果を発揮する PF を用いて数値実験を行い, データ同化の精度を検証するには最適なモデルである. しかしあくまで簡易モデルなため, このモデルにおいて集中豪雨などの顕著現象を再現するのは不可能である. また現実大気モデルを用いた場合に作成できるアンサンブル数は簡易モデルの場合よりずっと少ないため, 本実験は現実大気モデルで同様の数値実験を行う前の予備段階と言える.

Lorenz-96 モデルは積分方法として 4 次のルンゲ=クッタ法を用いた場合, 時間ステップ $\Delta t = 0.05$ で計算結果が安定する事が分かっており, また $\Delta t = 0.2$ が現実大気における 1 日の時間に相当するとされている (Lorenz 1996) . 更に, 自由度 N = 40, 外力項 F = 8 に設定するとカオス的な振る舞いになる事が確認されている (van Kekem, D. L. 2018) . 各グリッドに存在する変数 $x_i$ は非線形性に応じて値を変化させるが, 変数自体はあくまで数値であり物理的な意味を持たないため, 数値実験を行う者が変数 $x_i$ の変動幅から気温などの任意の物理量と見なす必要がある.

## 2.2 アンサンブルカルマンフィルタ

KF は,Kalman (1960) で考案された線形モデルに対するデータ同化のアルゴリズムであり, これは誤差の時間発展が線形成長で誤差の確率分布が正規分布である場合に, 推定誤差が最小となる解析値すなわち真値に最も近い最適解を求める事ができる. また EKF は KF を非線形モデルに対しても適用できる様に拡張させたアルゴリズムである. 本項では, 近藤 (2009) と三好 (2006) に基づき KF・EKF の説明と EnKF の導出を行う.

KF はある時刻の変数と誤差の時間発展から 1 時刻後の変数すなわち予報値と予報誤差を求めるプロセスと, カルマンゲインという観測値を予報値に同化する際の重みを予報誤差と観測誤差から求めるプロセスと, 観測値を予報値に同化して真値に最も近い最適解すなわち解析値と解析誤差を求める 3 つのプロセスからなる. これらのプロセスは大きく分けると 5 つの方程式からなる.

まず初めに数値予報モデルに初期値を与えて予報値を求める.

$$\mathbf{x}_t^f = \mathbf{M}\mathbf{x}_{t-1}^a \tag{2.3}$$

$\mathbf{x}_t^f$ は予報ベクトル,$\mathbf{M}$ はモデル行列,$\mathbf{x}_{t-1}^a$ は解析ベクトルである. また, 上付きの添え字の $a, f$ は, それぞれ解析 (analysis), 予報 (forecast) を表し, 下付きの添え字 $t$ は時刻を表す. 次に予

報誤差を求める.

$$\mathbf{P}_t^f = \mathbf{M}\mathbf{P}_{t-1}^a\mathbf{M}^\top + \mathbf{Q} \tag{2.4}$$

$\mathbf{P}_t^f$ は予報誤差共分散行列,$\mathbf{P}_{t-1}^a$ は解析誤差共分散行列,$\mathbf{Q}$ はモデル誤差共分散行列である. 更にカルマンゲインを求める.

$$\mathbf{K}_t = \mathbf{P}_t^f\mathbf{H}^\top\left(\mathbf{H}\mathbf{P}_t^f\mathbf{H}^\top + \mathbf{R}\right)^{-1} \tag{2.5}$$

$\mathbf{K}$ はカルマンゲイン,$\mathbf{R}$ は観測誤差共分散行列である. そしてカルマンゲインを用いて観測値を予報値に同化し, 解析誤差を求める.

$$\mathbf{P}_t^a = \left(\mathbf{I} - \mathbf{K}_t\mathbf{H}\right)\mathbf{P}_t^f \tag{2.6}$$

最後に解析値を求める.

$$\mathbf{x}_t^a = \mathbf{x}_t^f + \mathbf{K}_t\left(\mathbf{y}^o - \mathbf{H}\mathbf{x}_t^f\right) \tag{2.7}$$

$\mathbf{y}^o$ は観測値であり, これは真値に観測誤差を加えた値である. こうして求めた解析値を再び初期値として数値予報モデルに与える事で, 観測値によって予報値が修正される. またその情報がデータ同化サイクルにより将来にも伝わる事で 4 次元的にデータ同化が行われ, 数値予報モデルの予測精度が向上するという流れになっている.

ここまで誤差の時間発展が線形モデル $\mathbf{M}$ の場合について説明した. しかし, 現実大気モデルは非線形モデル $M$ であるが故にこのまま KF を適応する事はできない. そこで Jazwinski (1970) では, 非線形モデルを線形化することで KF を近似した EKF が考案された. 非線形モデル $M$ を $\mathbf{x}_0$ のまわりで線形化すると, 線形モデル $\mathbf{M}$ は,

$$\mathbf{M} = \left.\frac{\partial M}{\partial \mathbf{x}}\right|_{\mathbf{x}_t} \tag{2.8}$$

と表される. 式 (2.3) の線形モデル $\mathbf{M}$ を非線形モデル $M$ で置き換え, 式 (2.4) の線形モデル $\mathbf{M}$ には, 式 (2.8) を用いる. この様にする事で KF を非線形モデルに対しても適用する事が可能となる.

ただし線形アルゴリズムである KF を非線形系に適応すると誤差共分散行列を過小評価してしまう. 詳細は後述するが, それを防ぐ為に

$$\mathbf{P} \longrightarrow (1+\delta)\mathbf{P} \tag{2.9}$$

のように誤差共分散行列に対して 1 より僅かに大きい数をかける事で人為的に膨張させる必要がある. こうする事で誤差共分散行列の過小評価を防ぎ, データ同化サイクルを実行できる. この技術を共分散膨張と呼ぶ.

以上より,EKF を構成する式が揃ったため, 以下にそれらをまとめる. 予報のプロセスは以下の様になる.1 時刻前の状態変数 $\mathbf{x}_{t-1}^a$ から非線形モデル $M$ を用いて, 現在の状態変数 $\mathbf{x}^f$ を予報する.

$$\mathbf{x}_t^f = M\mathbf{x}_{t-1}^a \tag{2.10}$$

同様に 1 時刻前からの誤差の時間発展は次式で表される.

$$\mathbf{P}_t^f = \mathbf{M}\mathbf{P}_{t-1}^a\mathbf{M}^\top + \mathbf{Q} \tag{2.4}$$

解析のプロセスは以下の様になる. 解析誤差を最小にする様なカルマンゲイン $\mathbf{K}$ は次式で表される.

$$\mathbf{K}_t = \mathbf{P}_t^f\mathbf{H}^\top\left(\mathbf{H}\mathbf{P}_t^f\mathbf{H}^\top + \mathbf{R}\right)^{-1} \tag{2.5}$$

カルマンゲイン $\mathbf{K}$ を用いて, 予報 $\mathbf{P}^f$ を観測 $\mathbf{y}^o$ で修正する.

$$\mathbf{x}_t^a = \mathbf{x}_t^f + \mathbf{K}_t\left(\mathbf{y}^o - \mathbf{H}\mathbf{x}_t^f\right) \tag{2.7}$$

$\mathbf{K}$ を用いて予報誤差共分散行列 $\mathbf{P}^f$ を小さくし, 解析誤差 $\mathbf{P}^a$ を求める.

$$\mathbf{P}_t^a = \left(\mathbf{I} - \mathbf{K}_t\mathbf{H}\right)\mathbf{P}_t^f \tag{2.6}$$

となる.

KF が他のデータ同化手法と異なる点は, 時間変化する流れに依存した誤差共分散行列を陽に計算する点である. 最適内挿法や 3D-Var などでは予報誤差共分散行列には統計的な平均値を用いる. また計算効率を向上させるために予報誤差共分散行列を対角化して分散のみの行列に変形する. しかし KF は予報誤差共分散行列を直接計算し, かつ対角化も行わないため, 非対角成分の共分散も考慮したデータ同化が可能である. この共分散は異なる格子点間の相関および異なる変数間の相関に相当するため, 少ない観測値から多くの情報を予報値に同化

する事ができる. 例えば台風は海上で発生し, かつ通過する経路も異なるため, 直接台風の観測を行う事は難しい. しかし気象衛星は常時観測を行っているため,KF によって台風と相関のある衛星輝度温度の観測値を同化すれば, その予測精度を改善できる.KF のこの特徴は後述する EnKF でも同じである.

次に,EKF から EnKF が導出される過程を述べる.EKF を現実大気モデルに実装する場合, モデルの自由度が非常に大きい事が問題となる. 一般的な現実大気モデルの次元は $N = O(10^7 \sim 10^8)$ のオーダーであり, 予報誤差共分散行列 $\mathbf{P}^f$ は $N$ を一辺とする正方行列となるため, これを直接計算する事は計算機資源の観点から不可能である. そこで,EKF の予報誤差共分散行列 $\mathbf{P}'^f$ をアンサンブル予報を用いて近似した EnKF が Evensen (1994) によって提唱された. 本項では EKF の数式および行列を表す場合,$'$ を付けて EnKF のものと区別した.

まず, 状態変数の集合であるアンサンブルは,

$$\mathbf{X} = \left[\mathbf{x}^{(1)},\ \mathbf{x}^{(2)},\ \cdots,\mathbf{x}^{(m)}\right] \tag{2.11}$$

と表せる.$m$ はアンサンブルサイズであり, 一般的に大気の次元より大幅に小さい ($m \ll N$).1 時刻前のそれぞれの初期値 $\mathbf{X}^a_{t-1}$ からの予報は,

$$\mathbf{X}^f_t = M(\mathbf{X}^a_{t-1}) \tag{2.12}$$

と表すことができる. 予報誤差 $\mathbf{E}^f$ を各アンサンブルメンバーとアンサンブル平均 $\bar{\mathbf{X}}^f$ の差で表すと,

$$\begin{aligned} \delta\mathbf{x}^f_t &= \left[\mathbf{x}^{f(1)}_t - \bar{\mathbf{x}}^f_t,\ \mathbf{x}^{f(2)}_t - \bar{\mathbf{x}}^f_t,\ \cdots,\mathbf{x}^{f(m)}_t - \bar{\mathbf{x}}^f_t\right] \\ &= \mathbf{X}^f_t - \bar{\mathbf{X}}^f_t \end{aligned} \tag{2.13}$$

となる. なお,$\bar{\mathbf{X}}^f$ は真値ではない. 以上を用いると,EnKF の予報誤差共分散行列 $\mathbf{P}^f_t$ は,

$$\mathbf{P}^f_t = \frac{\delta\mathbf{x}^f_t\left(\delta\mathbf{x}^f_t\right)^\top}{m-1} \tag{2.14}$$

と表すことができる. 一方で,EKF の予報誤差共分散行列 $\mathbf{P}'^f_t$ は,

$$\mathbf{P}'^f_t = \frac{\delta\mathbf{x}'^f_t\left(\delta\mathbf{x}'^f_t\right)^\top}{N-1} \tag{2.15}$$

となる. ここで式 (2.14) と式 (2.15) を比較すると $m$ と $N$ が等しいので, アンサンブルサイズ $m$ は, モデルの次元 $N$ だけあれば十分である事が分かる. しかし, 計算機資源の観点からその様なアンサンブル予報は不可能である. そこで, 大気力学系の実質的な自由度は小さい (Dee 1995) という特徴を利用する. この特徴に基づいて,EKF の予報誤差共分散行列 $\mathbf{P}'^f$ を固有値分解すると固有値の多くはおよそ 0 になる. これは, 大気力学系の主成分分析を行うと多くの成分を 0 に近似できる, という事を示している. そこで, そのような固有値を無視すると,EKF の予報誤差共分散行列 $\mathbf{P}'^f$ は,

$$\mathbf{P}'^f = \frac{\delta\mathbf{x}_t^f\left(\delta\mathbf{x}_t^f\right)^\top}{m-1} \tag{2.16}$$

となり,EnKF の式と EKF の式が同値になる事が分かる. この結果に基づいて,EKF の方程式をアンサンブル予報誤差 $\delta\mathbf{x}_t^f$ を用いて EnKF の方程式に書き換えていく. 誤差の時間発展式 (2.4) は, モデル誤差 $\mathbf{Q}=0$ とすると,

$$\begin{aligned}\mathbf{P}_t'^f &= \mathbf{M}\mathbf{P}_{t-1}'^a\mathbf{M}^\top \\ &\simeq \frac{1}{m-1}\mathbf{M}\delta\mathbf{x}_{t-1}^a\left(\mathbf{M}\delta\mathbf{x}_{t-1}^a\right)^\top\end{aligned} \tag{2.17}$$

ここで,$\mathbf{M}\delta\mathbf{x}_{t-1}^a$ は,

$$\begin{aligned}\mathbf{M}\delta\mathbf{x}_{t-1}^a &= \left[\mathbf{M}\delta\mathbf{x}_{t-1}^{a(1)},\ \mathbf{M}\delta\mathbf{x}_{t-1}^{a(2)},\ \cdots,\ \mathbf{M}\delta\mathbf{x}_{t-1}^{a(m)}\right] \\ &\simeq \Big[M(\bar{\mathbf{x}}_{t-1}^a+\delta\mathbf{x}_{t-1}^{a(1)})-M(\bar{\mathbf{x}}_{t-1}^a),\ M(\bar{\mathbf{x}}_{t-1}^a+\delta\mathbf{x}_{t-1}^{a(2)})-M(\bar{\mathbf{x}}_{t-1}^a) \\ &\qquad \cdots,\ M(\bar{\mathbf{x}}_{t-1}^a+\delta\mathbf{x}_{t-1}^{a(m)})-M(\bar{\mathbf{x}}_{t-1}^a),\ \Big] \\ &= \left[M(\mathbf{x}_{t-1}^{a(1)})-\bar{\mathbf{x}}_t^f,\ M(\mathbf{x}_{t-1}^{a(2)})-\bar{\mathbf{x}}_t^f,\ \cdots,\ M(\mathbf{x}_{t-1}^{a(m)})-\bar{\mathbf{x}}_t^f\right]\end{aligned} \tag{2.18}$$

となり, この式は $m$ メンバーのアンサンブル予報と見なす事が可能であり, 非線形モデルを線形化する事なく, データ同化に利用できる. 以上より,EKF の予報誤差共分散行列 $\mathbf{P}_t'^f$ はアンサンブルメンバーを用いて,

$$\begin{aligned}\mathbf{P}_t'^f &= \mathbf{M}\mathbf{P}_{t-1}'^a\mathbf{M}^\top \\ &\simeq \frac{1}{m-1}\mathbf{M}\delta\mathbf{x}_{t-1}^a\left(\mathbf{M}\delta\mathbf{x}_{t-1}^a\right)^\top \\ &= \frac{\delta\mathbf{x}_t^f\left(\delta\mathbf{x}_t^f\right)^\top}{m-1} = \mathbf{P}_t^f\end{aligned} \tag{2.19}$$

の様に近似できる.EnKF は大気の実質的な自由度が大幅に小さい事を利用し, 効率的にアンサンブルメンバーをサンプリングする事で, 現実大気モデルの次元 $N$ より大幅に小さいサイズ $m$ のアンサンブルメンバーで, 誤差共分散行列を近似している. 次に, カルマンゲインについて考える. まず予報誤差共分散行列と観測誤差共分散行列からカルマンゲインを求める式 (2.5) に, 式 (2.19) を代入すると,

$$\begin{aligned}
\mathbf{K} &= \delta\mathbf{x}^f(\mathbf{H}\delta\mathbf{x}^f)^\top\left(\mathbf{H}\delta\mathbf{x}^f(\mathbf{H}\delta\mathbf{x}^f)^\top+(m-1)\mathbf{R}\right)^{-1} && (2.20)\\
&= \frac{1}{m-1}\delta\mathbf{x}_t^f\left(\mathbf{H}\delta\mathbf{x}_t^f\right)^\top\left[\frac{1}{m-1}\mathbf{H}\delta\mathbf{x}_t^f\left(\mathbf{H}\delta\mathbf{x}_t^f\right)^\top+(m-1)\mathbf{R}\right]^{-1}\\
&= \delta\mathbf{x}_t^f\left[(m-1)\mathbf{I}+\left(\mathbf{H}\delta\mathbf{x}_t^f\right)^\top\mathbf{R}^{-1}\mathbf{H}\delta\mathbf{x}_t^f\right]^{-1}\left(\mathbf{H}\delta\mathbf{x}_t^f\right)^\top\mathbf{R}^{-1} && (2.21)
\end{aligned}$$

の様に式変形できる. 式 (2.20) では, 全てのアンサンブルメンバーを保存できるだけのメモリは必要となるが,$N\times N$ 行列を直接コンピュータのメモリ上に保存する必要はなくなる. また, 式 (2.20) を式 (2.21) まで式変形すると,[ ] の中は $m\times m$ 行列となる. これにより逆行列の計算がアンサンブル数を一辺とした正方行列の計算を行うだけで済む様になる. 更に, 異なる観測値は互いに独立しており, 相関は無いと仮定すると観測誤差共分散行列 $\mathbf{R}$ は対角行列になるため, その逆行列を容易に計算できる様になる. また, 予報誤差共分散行列をその平方根 $\delta\mathbf{x}^f$ で近似すると, 以下の様に観測演算子 $\mathbf{H}$ を非線形観測演算子 $H$ として同じ形のまま式変形できる様になる.

$$\begin{aligned}
\mathbf{H}\delta\mathbf{x}_t^f &= \left[\mathbf{H}\delta\mathbf{x}_t^{f(1)},\ \mathbf{H}\delta\mathbf{x}_t^{f(2)},\ \cdots,\ \mathbf{H}\delta\mathbf{x}_t^{f(m)}\right]\\
&\simeq \Big[H(\bar{\mathbf{x}}_t^f+\delta\mathbf{x}_t^{f(1)})-H(\bar{\mathbf{x}}_t^f),\ H(\bar{\mathbf{x}}_t^f+\delta\mathbf{x}_t^{f(2)})-H(\bar{\mathbf{x}}_t^f)\\
&\qquad \cdots,\ H(\bar{\mathbf{x}}_t^f+\delta\mathbf{x}_t^{f(m)})-H(\bar{\mathbf{x}}_t^f),\ \Big]\\
&= \Big[H(\mathbf{x}_t^{f(1)})-H(\bar{\mathbf{x}}_t^f),\ H(\mathbf{x}_t^{f(2)})-H(\bar{\mathbf{x}}_t^f),\\
&\qquad \cdots,\ H(\mathbf{x}_t^{f(m)})-H(\bar{\mathbf{x}}_t^f)\Big] && (2.22)
\end{aligned}$$

最後に, 解析誤差共分散行列 $\mathbf{P}^a$ を計算する. 式変形を通じて, 予報誤差共分散行列 $\mathbf{P}^f$ をアンサンブル予報誤差 $\delta\mathbf{x}^f$ によって表現した様に, 解析誤差共分散行列 $\mathbf{P}^a$ も解析アンサンブル摂動 $\delta\mathbf{x}^a$ によって表現できるため, 直接 $\mathbf{P}^a$ を計算する必要はなくなる. 解析アンサンブル摂動 $\delta\mathbf{x}^a$ は, 解析誤差共分散行列 $\mathbf{P}^a$ の平方根であるため, これを求める事で解析誤差すなわちデータ同化の影響を反映した最適な摂動を作成する事が可能である. この解析アンサンブル

摂動 $\delta\mathbf{x}^a$ を計算し, 事後分布から 1 時刻後の事前分布を作成する過程をアンサンブルアップデートという.

アンサンブルアップデートの方法は大きく 2 種類に分けられる.1 つは PO-EnKF と呼ばれ, 観測値にも摂動を加える事で解析アンサンブル摂動を作成する. もう 1 つは Serial EnSRF で, こちらは解析誤差共分散行列を求める式 (2.27) を直接解く事で, 解析誤差共分散行列 $\mathbf{P}^a$ の平方根を計算し, 解析アンサンブル摂動を作成する. 次の項では,PO-EnKF と Serial EnSRF について述べていく.

## 2.3 PO-EnKF

PO-EnKF では, それぞれのアンサンブルメンバーに対して独立した解析予報サイクルを実行する. この際, モデル変数 $\mathbf{X}$ 以外に観測値にも摂動を加えてアンサンブルメンバーを作成する. 本項では, 西村 (2011) に基づき PO-EnKF の導出を行う.

$$\mathbf{Y}^o = [\bar{\mathbf{y}}^o + \delta\mathbf{y}^o] \tag{2.23}$$

ここで, 観測値 $\bar{\mathbf{y}}^o$ には既に観測誤差が含まれており, これに摂動 $\delta\mathbf{y}^o$ を加えた $\mathbf{Y}^o$ を解析予報サイクルで利用する. すると,PO-EnKF における予報値を観測値で修正する式, すなわち同化方程式は,

$$\begin{aligned} \mathbf{X}^a &= \mathbf{X}^f + \mathbf{K}\left(\mathbf{Y}^o - \mathbf{H}\mathbf{X}^f\right) \\ \bar{\mathbf{X}}^a + \delta\mathbf{X}^a &= \bar{\mathbf{X}}^f + \delta\mathbf{X}^f + \mathbf{K}\left[\bar{\mathbf{Y}}^o + \delta\mathbf{Y}^o - \mathbf{H}\left(\bar{\mathbf{X}}^f + \delta\mathbf{X}^f\right)\right] \end{aligned} \tag{2.24}$$

となる. よって, それぞれの摂動に対しても

$$\delta\mathbf{X}^a = \delta\mathbf{X}^f + \mathbf{K}\left(\delta\mathbf{Y}^o - \mathbf{H}\delta\mathbf{X}^f\right) \tag{2.25}$$

が成り立つ. ここで, 観測に摂動が無い場合を考えると,

$$\begin{aligned} \delta\mathbf{X}^a &= \delta\mathbf{X}^f - \mathbf{K}\mathbf{H}\delta\mathbf{X}^f \\ \delta\mathbf{X}^a &= [\mathbf{I} - \mathbf{K}\mathbf{H}]\,\delta\mathbf{X}^f \end{aligned} \tag{2.26}$$

となる. この式 (2.26) は, それぞれのアンサンブルメンバーに対して独立した解析予報サイクルを実行する際に, 全てに対して同じ観測値を同化する事を示している. この式 (2.26) から, 解析誤差共分散行列 $\mathbf{P}^a$ を求めると,

$$\begin{aligned}\mathbf{P}^a &= \frac{1}{m-1}\delta\mathbf{X}^a\left[\delta\mathbf{X}^f\right]^\top \\ &= \frac{1}{m-1}\left(\mathbf{I}-\mathbf{KH}\right)\delta\mathbf{X}^f\left[\left(\mathbf{I}-\mathbf{KH}\right)\delta\mathbf{X}^f\right]^\top \\ &= \left[\mathbf{I}-\mathbf{KH}\right]\mathbf{P}^f\left[\mathbf{I}-\mathbf{KH}\right]^\top \end{aligned} \tag{2.27}$$

となる. しかし, カルマンフィルタにおける解析誤差共分散行列 $\mathbf{P}^a$ は

$$\begin{aligned}\mathbf{P}_t^a &= \left\langle\delta\mathbf{x}_t^a\left(\delta\mathbf{x}_t^a\right)^\top\right\rangle \\ &= \left\langle\left(\left(\mathbf{I}-\mathbf{K}_t\mathbf{H}\right)\delta\mathbf{x}_t^f+\mathbf{K}_t\delta\mathbf{y}^o\right)\left(\left(\mathbf{I}-\mathbf{K}_t\mathbf{H}\right)\delta\mathbf{x}_t^f+\mathbf{K}_t\delta\mathbf{y}^o\right)^\top\right\rangle \\ &= \left(\mathbf{I}-\mathbf{K}_t\mathbf{H}\right)\mathbf{P}_t^f\left(\mathbf{I}-\mathbf{K}_t\mathbf{H}\right)^\top+\mathbf{K}_t\mathbf{R}\mathbf{K}^\top+cross \end{aligned} \tag{2.28}$$

となる.$cross$ は予報誤差と観測誤差のクロスタームであり, 予報と観測は独立で相関が無いため, 値は 0 になる. 式 (2.27) と式 (2.28) を比較すると PO-EnKF は KF に比べて $\mathbf{K}_t\mathbf{R}\mathbf{K}^\top$ が欠けており, これは観測誤差を考慮しなかった事で解析誤差が過小評価されている事を示している. そこで, この問題を回避する為に観測に摂動を加えて式 (2.23) の様に観測データをアンサンブル化するのである.

以上が PO-EnKF のアルゴリズムであり, このデータ同化手法はそれぞれのアンサンブルメンバーに対して独立した解析を行うため, 一度に全てのアンサンブルメンバーをメモリに格納する必要がなく, 計算コストが少ないという特徴がある. しかし, 観測に摂動を与えるという事は, 新たなサンプリングエラーを導入するという事に他ならず, アンサンブル数が少ない場合は観測に摂動を与えない Serial EnSRF の方が, データ同化の解析精度が優れるとされている (Whitaker and Hamill 2002).

## 2.4 Serial EnSRF

SRF は解析誤差共分散行列から解析誤差を求める事でアンサンブルアップデートを行う手法である.SRF は PO-EnKF と異なり, それぞれのアンサンブルメンバーに対してではなく, ア

ンサンブル平均に対して解析を行う. また, 解析アンサンブルは予報アンサンブルに変換行列をかけて求める.SRF では解析誤差を解析誤差共分散行列の平方根を解く事で求めるため, 解析値に含まれる解析誤差を反映した適切なアンサンブルを作成できる. 本項では, 西村 (2011) に基づき Serial EnSRF について導出を行う.

SRF のアンサンブルアップデートは, 変換行列 $\mathbf{T}$ を用いて行う.

$$\delta\mathbf{x}^a = \delta\mathbf{x}^f\mathbf{T} \tag{2.29}$$

特に Serial EnSRF は,SRF の中で最もシンプルなデータ同化手法の 1 つである.Serial EnSRF では,

$$\delta\mathbf{x}^a = \left(\mathbf{I} - \tilde{\mathbf{K}}\mathbf{H}\right)\delta\mathbf{x}^f \tag{2.30}$$

という形でアンサンブルアップデートを行う. 上述した変換行列 $\mathbf{T}$ は, ここでは $\mathbf{I} - \tilde{\mathbf{K}}\mathbf{H}$ に相当する.

式 (2.30) が式 (2.28) を満たすには

$$\begin{aligned}
\mathbf{P}^a &= \left\langle \delta\mathbf{x}_t^a\left(\delta\mathbf{x}_t^a\right)^\top \right\rangle \\
&= \left(\mathbf{I} - \tilde{\mathbf{K}}\mathbf{H}\right)\delta\mathbf{x}^f(\delta\mathbf{x}^f)^\top\left(\mathbf{I} - \tilde{\mathbf{K}}\mathbf{H}\right)^\top \\
&= \left(\mathbf{I} - \tilde{\mathbf{K}}\mathbf{H}\right)\mathbf{P}^f\left(\mathbf{I} - \tilde{\mathbf{K}}\mathbf{H}\right)^\top \\
\left(\mathbf{I} - \tilde{\mathbf{K}}\mathbf{H}\right)\mathbf{P}^f\left(\mathbf{I} - \tilde{\mathbf{K}}\mathbf{H}\right)^\top &= \left(\mathbf{I} - \mathbf{K}\mathbf{H}\right)\mathbf{P}^f \qquad (2.31)
\end{aligned}$$

となる必要がある.Andrew (1986) によると, 式 (2.31) から求まるカルマンゲイン $\mathbf{K}$ は

$$\mathbf{K} = \delta\mathbf{x}^f\left(\mathbf{H}\delta\mathbf{x}^f\right)^\top\left[\mathbf{H}\delta\mathbf{x}^f\left(\mathbf{H}\delta\mathbf{x}^f\right)^\top + (m-1)\mathbf{R}\right]^{-1} \tag{2.20}$$

となる.

Serial EnERF では, 観測がそれぞれ独立であると仮定し, 式 (2.21) のカルマンゲインを用いて観測を逐次的に同化する. したがって, 観測演算子 $\mathbf{H}$, カルマンゲイン $\mathbf{K}$ は $n$ 次元ベクトルとなり, 観測誤差共分散行列 $\mathbf{R}$ と $\mathbf{H}\mathbf{P}^f\mathbf{H}^\top$ は $1*1$ 行列なので実質的にスカラーになる. それにより, 式 (2.21) は行列の演算から方程式へ簡略化され,$\tilde{\mathbf{K}}$ は,$\tilde{\mathbf{K}} = \alpha\mathbf{K}$ という形になる. こ

の時,$\alpha$ は (Whitaker and Hamill, 2002) によると,

$$\alpha = \left(1 + \sqrt{\frac{\mathbf{R}}{\mathbf{HP}^f\mathbf{H}^\top + \mathbf{R}}}\right)^{-1} \tag{2.32}$$

となる.

このデータ同化手法では, 逐次的に観測を 1 つずつ同化していくため, 全ての観測を同化するまで式 (2.21) を繰り返し計算していく必要がある. この様な性質から,Serial EnSRF は並列化する事が非常に困難であり, データ同化の解析精度は高いものの計算効率は低い. 並列化は, 大量のメモリを使用する現実大気モデルに EnFK を実装する上で, 使用するメモリを分散させ, 計算速度を向上させる為には欠かせない技術である. そこで, 次項では EnKF を実装する上で, 計算負荷を軽減し, かつデータ同化の解析精度を向上させる為に必要となる局所化の概念について記述する.

## 2.5 局所化

局所化は現実大気モデルに EnKF を実装する上で限られたアンサンブル数を原因とするサンプリングエラーを抑える為に欠かせない技術である. 局所化では, あるモデル格子点から遠く離れた観測はモデル格子点に与える影響が殆ど無いものと仮定し, あるモデル格子点から一定の距離の内側に存在する観測のみを同化する. 例えば太平洋における気温の観測を同化する事で東京における気温の予報精度が向上するというのは考え難く, 両者に誤差相関がないという仮定は尤もらしい. しかし, 実際にはある地点の大気現象が遠く離れた地点に全く影響を与えない訳ではなく, 現実大気にはテレコネクションという概念が存在する. 例えば, 太平洋における東西方向の海面水温偏差がエルニーニョやラニーニャという現象を通じて, テレコネクションにより日本の気候に影響を与える事は一般的にもよく知られた事実である. したがって, 厳密には遠く離れた地点の観測を同化する事によって得られるインパクトも存在する. 実際,Kondo and Miyoshi (2016) では簡易大気大循環モデル (SPEEDY: Simplified Parameterizations primitivE-Equation DYnamics) を用いた 10240 メンバーによる局所化を無くした実験を行い, 遠く離れた地点の観測が与えるインパクトを調べている. 本項では, 近

藤 (2009) に基づき, 局所化について導出を行う.

EnKF は限られたアンサンブルメンバーで実行すると, サンプリングエラーによりデータ同化の解析精度が低下する. このサンプリングエラーによる影響は全球上のどの地点であっても一定である. 一方で, 観測を同化する事で得られるモデル格子点に対する影響は, 上述した様に観測点とモデル格子点の距離に反比例する. したがって, 観測点がモデル格子点から離れるにつれて, サンプリングエラーによる影響の割合は相対的に大きくなる. これは, 遠く離れた地点の観測を同化すると, サンプリングエラーによってデータ同化の解析精度が低下する事を意味している. アンサンブル数を増やせばサンプリングエラーは少なくなり,EnKF と KF は同値になる. しかし, 実際には計算機資源は限られており, かつ現実大気モデルは巨大な次元を有するため, 作成できるアンサンブル数は数十〜数百程度が限界である. したがって, 現実大気モデルにおいて EnKF の解析精度を十分保つ為には, 局所化を行う事で同化する観測範囲を制限し, 遠く離れた地点の観測に起因するサンプリングエラーを軽減する必要がある.

局所化は, 予報誤差共分散行列に局所化関数をかけて, 共分散すなわち非対角成分の値を 0 にする事で, 遠く離れた地点との誤差相関を無くす. 局所化関数はどの様な関数でも良いが, 一般的にはガウス関数を用いる. ただし, 純粋なガウス関数では中央値からどれほど離れた値であっても厳密には 0 にならないため, 何かしらの閾値を設定して, その閾値を超えると値が 0 になる様にする. 本研究で用いた局所化関数もガウス関数であり,$r = 2\sqrt{10/3}$ より遠くの格子点で観測の重みが 0 になる様にしている (Gaspari and Cohn 1999). なお本研究では局所化スケールを 1〜10 の範囲でチューニングしているため, 局所化スケールが 1 であれば周囲およそ 3.65 格子点分の範囲にある観測のみ同化する事になる.

上述した様に, 現実大気モデルにおいて予報誤差共分散行列 $\mathbf{P}^f$ を直接計算する事はできない. したがって, 本研究における EnKF ではカルマンゲイン $\mathbf{K}$ に局所化関数 $L(r)$ をかけて局所化を行う. これを R-Localization と言い, 予報誤差共分散行列に局所化関数をかけて直接計

算する B-Localization を近似したものに相当する.

$$
\begin{aligned}
\mathbf{K} &= \left(L(r)\mathbf{P}^f\right)\mathbf{H}^\top\left(\mathbf{H}\left(L(r)\mathbf{P}^f\right)\mathbf{H}^\top + \mathbf{R}\right)^{-1} \\
&\simeq L(r)\left(\mathbf{P}^f\mathbf{H}^\top\left(\mathbf{H}\mathbf{P}^f\mathbf{H}^\top + \mathbf{R}\right)^{-1}\right) \qquad (2.33)
\end{aligned}
$$

一方,PF では予報誤差共分散行列やカルマンゲインを計算しないため,EnKF とは異なる方法で局所化を行う. 一定の距離すなわち局所化半径より遠い格子点では観測誤差の値を無限大にする事で, 観測が予報に同化されない様にする. これは言い換えると,EnKF では全格子点で局所化を行うのに対し,PF では各格子点で局所化を行うという事になる.

## 2.6 共分散膨張

EKF では, 非線形モデルを線形化する際に, 予報誤差共分散行列が過小評価される. これは, 非線形方程式を解く為に摂動法を用いて線形化する際は, 状態変数を基本場と擾乱に分けて擾乱を無視するが, それによって状態変数の値が線形化を行う前より小さくなり, 擾乱が無くなった分だけ誤差共分散が小さくなる事が一因として考えられる. 一方の EnKF では, アンサンブル数が少ないために, 予報誤差共分散行列が過小評価される.EnKF では大気力学系の自由度が実質的には小さい事を利用して, 大気の次元 $N$ より大幅に少ないアンサンブル数 $m$ で誤差共分散行列を近似している. しかし, これは大気力学系の主成分分析を行った際に得られる小さい成分を無視するという事であり, その分だけ誤差共分散が小さくなる.

予報誤差共分散行列が過小評価された状態で解析予報サイクルを繰り返すと,EKF や EnKF は予報を過大評価して観測を同化しなくなるため, やがてフィルタ発散に陥る. この事は, カルマンゲインを求める式 (2.5) と同化方程式 (2.6) からも分かる. 予報誤差共分散行列 $\mathbf{P}^f$ が小さくなる事で観測に対する重み $\mathbf{K}$ が小さくなり, 観測の情報が解析に反映され難くなる事で, やがてモデルが現実大気から離れて独り歩きを始めてしまう. この様になるのは, 線形アルゴリズムの KF を非線形モデルに適用する際に,KF の理論が満たされない部分が生じる為だと考えられている (Miyoshi 2006).

そこでフィルタ発散を防ぐ為に導入された, 予報誤差共分散行列を人為的に大きくする共分散膨張という技術について, 本項では西村 (2011) に基づき導出を行う.

$$\mathbf{P}^f \longrightarrow (1+\delta)\mathbf{P}^f \tag{2.34}$$

ここで, 膨張係数 $\delta$ は 0 より僅かに大きい数である. この様に, 予報誤差共分散行列 $\mathbf{P}^f$ に 1 より僅かに大きい数をかける事で, 予報誤差共分散行列を膨張させ, 予報誤差の過小評価を防ぐ. また, 誤差共分散の過小評価をモデル誤差に相当するものとして表現する事も可能である.

$$(1+\delta)\mathbf{P}^f = \mathbf{P}^f + \mathbf{Q} \tag{2.35}$$

しかし, この $\mathbf{Q}$ は必ずしもモデル誤差による影響のみを表現している訳ではない. したがって, 共分散膨張はあくまで EKF や EnKF が発散しない様にする為の技術に過ぎず, 共分散膨張を必要とする理由が理論的に明らかになっている訳ではない.

EnKF において, 共分散膨張を行う方法は複数あるが, ここではその中から代表的な方法を 2 つ紹介する.1 つは誤差共分散を直接膨張させず, 誤差共分散を計算する前の予報誤差を膨張させる,multiplicative inflation (Pham 1998) という方法である.

$$\delta\mathbf{x}^f \longrightarrow (1+\delta)\delta\mathbf{x}^f \tag{2.36}$$

こうする事で予報誤差共分散行列 $\mathbf{P}^f$ は $(1+\delta)^2$ 倍になる.

もう 1 つは, 解析アンサンブルに対して統計的に計算した解析誤差共分散行列を反映させる様なランダムな摂動 $\eta$ を加える,additive inflation (Corazza et al. 2002) という方法である.

$$\mathbf{E}^a \longrightarrow \mathbf{E}^a + \eta \tag{2.37}$$

ただし, アンサンブル数が少ない場合, 摂動 $\eta$ から計算される統計的な解析誤差共分散行列には, サンプリングエラーが含まれる. そこで, ランダムな摂動を予報誤差ではなく解析誤差に加える事で, 予報プロセスを通じてランダム摂動 $\eta$ にモデルをかける.

$$\begin{aligned}\mathbf{E}^f &= \mathbf{M}\left(\mathbf{E}^a + \eta\right) \\ &= \mathbf{M}\mathbf{E}^a + \mathbf{M}\eta \end{aligned} \tag{2.38}$$

すると, 科学的な意味を持たないランダム摂動 $\eta$ が, モデルの影響を受けた $\mathbf{M}\eta$ になる. こうする事で, 日々変動する誤差成分が取り出され, サンプリングエラーが減少するため, ランダ

ム摂動がより適切な誤差になる.

Miyoshi et al. (2005) では,40 変数の Lorenz-96 モデルを使って, 過去から予報時刻までに最も成長した誤差ベクトル (BV: Bred Vector) と, ランダム摂動を足してモデルをかけた BV を比較した. その結果, 予報プロセスを通じてモデルによる力学的拘束を受けた BV の方が, 流れに依存したより適切な誤差成分を扱う事が可能であり, かつ外力が加えられない場合は, ランダム摂動が成長しない事が明らかになった. ここで紹介したもの以外に, 共分散膨張の方法は Zhang et al. (2004) により Relaxation to prior perturbation や, Whitaker and Hamill (2012) により Relaxation to prior spread というものが考案されており, 近年ではこれらの方法が主流になっている.

KF は線形アルゴリズムであるため, カオスの影響が現れない線形モデルにおいては過去の情報を未来まで反映させ続ける事ができる. しかし, 現実大気モデルは非線形モデルであるため, カオスの影響によって過去の情報が反映されなくなり, 誤差共分散が過小評価される様になる. つまり, 共分散膨張は言い換えると, 非線形モデルに過去の情報を未来まで反映させ続ける為の技術という事になる.

## 2.7 粒子フィルタ

PF は Kitagawa, G. (1993) で提唱されたデータ同化のアルゴリズムである. これまで述べてきた EnKF は非線形モデルに対しても適用できるデータ同化手法であり, 多くの点で優れているため, 現業数値予報モデルや研究の分野で幅広く利用されている. しかし, 誤差の時間発展が線形で, かつ誤差の確率分布がガウス分布である場合に推定誤差が最小となる解析値を求める. そのため, 強い非線形性によって予報誤差が非ガウス分布になる顕著現象に対しては有効に機能せず, 例えば集中豪雨の予測を行う際は解析の精度が低下すると考えられている. そこで本項では樋口 (2011) に基づき, その様な仮定を一切せず, より一般的なモデルに対しても適用できる PF について導出を行う.

PF は EnKF と同様に状態の確率分布を多数のサンプルすなわち粒子で表現する, アンサン

ブル近似を用いたアルゴリズムである. したがって, モデルを用いてある時刻の状態と誤差の時間発展から 1 時刻後の変数を求める. つまり, 予報のプロセスまでは EnKF と同様である. しかし, 解析のプロセスが EnKF と大きく異なり,EnKF では KF のアルゴリズムを踏襲して様々な仮定を経た上で解析を行うが,PF ではそれらの仮定を一切せず粒子のみ用いて確率分布を表現する. これにより, 観測とモデルが非線形関係の場合や, 誤差の確率分布が非ガウス分布になる様な複雑なモデルに対しても適用する事が可能になっている. ここでは,KF の導出および解説を行った時と同様に, まず 4 つの方程式を紹介する事で PF のプロセスを大まかに把握し, それからより詳細に関して導出する事で PF のアルゴリズムを解説する.

まず，状態変数 $\mathbf{x}^a_{t-1}$ に予報モデル $M$ を作用させる事で，1 時刻前の解析値 (初期値) から現在の予報値を求める．

$$\mathbf{x}^f_t = M(\mathbf{x}^a_{t-1}) \tag{2.12}$$

ここまでは EnKF と同様である. 次に予報値と観測値の誤差について考える.EnKF では予報誤差共分散行列を, 解析誤差共分散行列と予報モデルおよびモデル誤差共分散行列から求めていた. しかし,PF では尤度という概念を用いて予報値の尤もらしさを求める.

$$p(\mathbf{y}_t \mid \mathbf{x}^i_{t|t-1}) = \frac{1}{\sqrt{(2\pi)^l|\mathbf{R}_t|}} \exp\left[-\frac{1}{2}\left(\mathbf{y}_t - \mathbf{H}_t(\mathbf{x}^i_{t|t-1})\right)^\top \mathbf{R}_t^{-1}\left(\mathbf{y}_t - \mathbf{H}_t(\mathbf{x}^i_{t|t-1})\right)\right] \tag{2.39}$$

観測誤差が平均 0, 観測誤差分散行列 $\mathbf{R}$ のガウス分布に従うと仮定すると,$\mathbf{x}^f(i)$ の尤度 $p(\mathbf{y}_t \mid \mathbf{x}^i_t)$ は以上の様に求められる. ここで $\mathbf{x}_{t|t-1}$ は予報値に相当し, $\mathbf{y}$ は観測値, $\mathbf{H}$ は観測演算子, 上付き添字 $i$ は粒子の番号, 下付き添字 $t$ は時刻, 上付き添字 $l$ は観測値 $\mathbf{y}_t$ の次元を示す. 尤度は観測値に対する予報値の尤もらしさである. 尤度を計算する関数には一般的にガウス関数が用いられるが, 他により適切な関数が存在すれば, その関数を用いる事ができる. 次に, 各粒子の重みを求める.EnKF では予報誤差共分散行列と観測誤差共分散行列からカルマンゲイン, すなわち観測の重みを求めていた. しかし,PF では粒子間の重みを尤度を用いて表現する.

$$w^i_t = \frac{p(\mathbf{y}_t \mid \mathbf{x}^i_t)}{\sum_{i=1}^N p(\mathbf{y}_t \mid \mathbf{x}^i_t)} \tag{2.40}$$

ある粒子の重み $w_i$ は, 尤度の総和 $\sum_{i=1}^N p(\mathbf{y}_t \mid \mathbf{x}^i_t)$ に対する, ある粒子の重みの割合によって表される. こうして求められた重みに応じて粒子を再配置する事で, より確からしい確率分布, すなわち解析値を得る事ができる.EnKF では予報アンサンブルとカルマンゲインから解

析アンサンブルを求めていた. 同様に,PF でも予報アンサンブルと重みから解析アンサンブルを求める.

$$p(\mathbf{x}_t \mid \mathbf{y}_{1:t}) = \sum_{i=1}^{N} w_t^i \delta(\mathbf{x}_t - \mathbf{x}_{t|t-1}^i) \tag{2.41}$$

ここで, $\delta$ はディラックのデルタ関数である.PF では尤度に応じた粒子の再配置, すなわちリサンプリングによって解析値を求める. この際, 尤度の低い粒子を破棄して, 尤度の高い粒子を複製する. この時, 複製される粒子を全く同じものにするか, それとも摂動を与えて僅かに異なる粒子にするかでリサンプリングの方法が異なってくるが, 本研究では全く同じ粒子を複製する事にした. したがって, 解析予報サイクルを繰り返す内に同じ粒子が複製されていき, アンサンブルすなわち粒子全体に占める同一の粒子が増加していく. すなわち有効粒子数が低下する. 解析値を初期値として再び予報モデルに与える際に, 初期値に乱数を加えたり,stochastic モデルを使うことにより各粒子はある程度散らばるが, それでも重みの大きい数少ない特定の粒子が複製されていく事で, 粒子全体が似た構造になってしまう. こうなると, 本来の広がりを持った確率分布が表現できなくなり, モデルが現実から離れて独り歩きを始めてしまう. この現象はアンサンブルの退化と呼ばれ, リサンプリングによって解析値を得る PF に特有の現象である.

アンサンブルの退化を防ぐ為には粒子数を増やすか, リサンプリングの方法を工夫する必要がある. 前者については, 粒子数が増えると計算コストは増加し, 現実大気モデルともなれば作成できる粒子数は数百〜数千が限界であるため, 計算機資源の観点から現実的とは言えない. 後者については, 様々な方法が考案されているが, ここではその中でも代表的な Probabilistic resampling と Stochastic universal sampling について紹介する.Probabilistic resampling では一様乱数に基づいて複製する粒子を抽出する. 各粒子の重みは異なっているため, 乱数に基づいて抽出すれば重みが大きい粒子がより多く選択され, 解析値の確率分布は重みを反映したものになる. 粒子が無限大に存在すれば一様に乱数は生成されるが, 粒子数が少なければ抽出される粒子にサンプリングエラーによる偏りが生じてしまうという欠点があり, 効率的ではない. 一方,Stochastic universal sampling では一様乱数を用いず, 区間 [0,1] を粒子数で等間隔に区切り, 区切った位置にある粒子を選択する. 抽出の際に乱数を用いないため, 粒子数が少ない場合も偏りが生じ難いという点で優れている.

リサンプリングの過程では, 予報アンサンブル $\delta\mathbf{x}^f$ に変換行列 $\mathbf{T}$ をかけて, 解析アンサンブル $\delta\mathbf{x}^a$ を得るという操作を行う.

$$\delta\mathbf{x}^a = \delta\mathbf{x}^f\mathbf{T} \tag{2.42}$$

式を見ると分かる様に, この式は SRF のアンサンブルアップデートの式と同じである. ここから,PF ではリサンプリングが EnKF におけるアンサンブルアップデートに相当する, という事が分かる. 変換行列 $\mathbf{T}$ は, 抽出された粒子を 1, 抽出されなかった粒子を 0 とする要素から構成されている.

更に, 本研究では同じ粒子の複製を緩和するため, 局所化を用いている. 格子点ごとに局所化範囲内の観測を使って尤度を求めリサンプリングを行うことにより, 各格子点では同じ値を持つ粒子が複製されても, 全格子点で同じ値を持つ粒子はほとんど存在しないと想定している.

この様に EnKF と PF のアルゴリズムが大きく異なるのは,PF が EnKF を踏襲したガウス仮定を置いていない為であり, それ故に EnKF には存在するが PF には存在しない概念も多い. たとえば, 共分散膨張は非線形モデルを線形化する事に起因する誤差共分散の過小評価を防ぎ, かつ少ないアンサンブル数を補う為のものだが,PF では共分散行列を陽に扱わないため, 共分散膨張という技術が存在しない.

ただし、PF にも共分散膨張に近い技術は存在し、それは以下の数式によって説明される。

$$w^i = \tau w^i + (1-\tau)/N, \quad (0 \leqq \tau \leqq 1) \tag{2.43}$$

この技術は厳密には共分散膨張と異なるものだが、ここでは便宜上、PF に共分散膨張として $\tau$ を膨張係数と呼ぶ。膨張係数の値を 1 にすると、ある粒子の重み $w^i$ はそのままの値となる。しかし、膨張係数の値を 0 にするとある粒子の重み $w^i$ は、アンサンブル数 $N$ の逆数となる。これは、膨張係数の値を 1 に近づけると特定の粒子の重みが相対的に大きくなり、粒子全体がその特定の粒子と似た構造になる事を示している。対して、膨張係数の値を 0 に近づけると全ての粒子の重みが近い値になり、粒子の多様性が維持される。このチューニングパラメータを適切な値にすれば、アンサンブルの退化を防ぐと同時に、粒子全体が真値を

より効率的に捉える様になり、データ同化の解析精度が向上する。

共分散膨張と同様に局所化も EnKF と PF ではアルゴリズムが大きく異なっており,PO-EnKF では予報誤差共分散行列に,Serial EnSRF ではカルマンゲインに, それぞれ局所化関数をかける事で表現していたが,PF には予報誤差共分散行列もカルマンゲインも存在しないため, 異なる方法で局所化を実現する.PF において, 局所化は観測空間で行われ, 格子点ごとに予報アンサンブルに変換行列 $\mathbf{T}$ をかけて解析アンサンブルを得る過程で行われる. つまり,PF では格子点ごとに局所化スケールに応じて異なる変換行列 $\mathbf{T}$ を求める事となる.

上述した様に,PF はガウス仮定を置かない為に EnKF よりも一般的なモデルに適用できるが, その分多くの粒子数を必要とする. また、データ同化に伴ってアンサンブルの退化という新たな問題が生じうる. 現実大気モデルに PF を実装する為にはこれらの問題を克服する必要があり, その為の研究が近年では盛んである. そこで本研究でも, 膨張係数や局所化スケールなどのチューニングパラメータを変化させて数値実験を繰り返す事で, どの様な条件下であれば PF が機能し, かつ EnKF よりもデータ同化の解析精度が向上するのか明らかにしていく.

# 第3章　双子実験

## 3.1　最適な解析条件の推定

本研究では,Lorenz-96 モデルに PO-EnKF・Serial EnSRF・PF を実装してデータ同化システムを構築し, 数値実験を行った. 真値は,Lorenz-96 モデルの長期ランによって作成した. この時,Lorenz-96 モデルの自由度 $i = 40$, 外力項 $F = 8$, 時間間隔 $\Delta t = 0.01$ にした. 観測値は, 真値に正規乱数の観測誤差を加えて作成した. 観測誤差の解析 RMSE は 1.0, 時間間隔 $\Delta t = 0.05$, データ同化のサイクル数は 2920 である.Lorenz-96 モデルでは $\Delta t = 0.2$ が現実大気の 1 日に相当するため,2 年分のデータ同化サイクルを実行した事になる. これは, データ同化が完全に収束するまでに 1 年間あれば充分であると考え, 最初の 1 年間をスピンアップの期間とした為である. また, 観測密度については全格子点で観測を行っている.

この実験では, 全期間におけるモデル領域全体の解析 RMSE の等値線図を作成した. 縦軸に膨張係数, 横軸に局所化スケールを取った解析 RMSE の等値線図を作成する事で, それらチューニングパラメータを変化させた場合, どの様な条件下ならばデータ同化の解析精度が最も良くなるか確認した. なお、膨張係数と局所化スケールに単位は伴わない。本実験における局所化関数はガウス関数であり,$r = 2\sqrt{10/3}$ より遠くの格子点で観測の重みが 0 になる様にしている. 局所化スケールは 1〜10 まで 1 ずつ, 膨張係数は 0.01〜0.1 まで 0.01 ずつ変化させた. 上述した様に,EnKF には共分散膨張や局所化といった概念が存在するが,PF には共分散膨張という概念が厳密には存在しないため,EnKF と同様の等値線図は作成していない. 以下に,PO-EnKF でアンサンブル数を 16 と 1000,Serial EnSRF でアンサンブル数を 16 と 1000 にした場合の解析 RMSE の等値線図を示す.

なお、解析 RMSE は以下の数式に基づいて求めている。

$$RMSE_{analysis} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\mathbf{x}^a - \mathbf{x}^t)^2} \tag{3.1}$$

解析 RMSE は、真値と解析値の二乗平均平方根誤差である.

### 3.1.1 結果

図 1 は, データ同化手法を PO-EnKF, アンサンブル数を 16 にした時の真値と解析値の解析誤差の等値線図である. 膨張係数を 0.05〜0.10, 局所化スケールを 4〜6 にした時の解析 RMSE が 0.24 と最も小さく、データ同化の解析精度が高い. 一方, 膨張係数を 0.01, 局所化スケールを 2 にした時の解析 RMSE と, 膨張係数を 0.10, 局所化スケールを 7 にした時の解析 RMSE は等値線によって結ばれており, この等値線を境に解析 RMSE の値が大きく異なっている. 境界線の右側では解析 RMSE が左側のおよそ 10 倍になっており, 観測誤差の解析 RMSE である 1.0 を超えているため、フィルタ発散に陥っている. また、境界線に注目すると膨張係数と局所化スケールの線形関係の様なものが見られ, 膨張係数が大きくなる事に比例して局所化スケールも大きくなる様に境界線が存在している.

図 2 は, データ同化手法を PO-EnKF, アンサンブル数を 1000 にした時の真値と解析値の解析誤差の等値線図である. 膨張係数を 0.01〜0.03, 局所化スケールを 6〜10 にした時の解析 RMSE が 0.2 以下と最も小さく、データ同化の解析精度が高い. 全体的な傾向としては、膨張係数が小さく、かつ局所化スケールが大きくなるほど解析精度が向上している。

図 3 は, データ同化手法を Serial EnSRF, アンサンブル数を 16 にした時の真値と解析値の解析誤差の等値線図である. 膨張係数を 0.01〜0.05, 局所化スケールを 7〜9 にした時の解析 RMSE が 0.19 と最も小さく、データ同化の解析精度が高い. 全体的な傾向としては、局所化スケールが大きくなるほど解析精度が向上している。

図 4 は, データ同化手法を Serial EnSRF, アンサンブル数を 1000 にした時の真値と解析値の解析誤差の等値線図である. 膨張係数を 0.01〜0.03, 局所化スケールを 8〜10 にした時の解

析 RMSE が 0.23 と最も小さく、データ同化の解析精度が高い. 全体的な傾向としては、膨張係数が小さく、かつ局所化スケールが大きくなるほど解析精度が向上している。

## 3.2 解析精度の時間変化

モデル領域全体の解析 RMSE と解析 SPREAD, すなわち解析アンサンブルの広がりの時系列図を作成した. 時系列図を作成する事でデータ同化システムが正常に機能し続けているか, また解析アンサンブルが真値を捉え続けているか確認した. なお、解析 RMSE と解析 SPREAD に単位は伴わず、これらは同じ次元であるため縦軸を共有しており、いずれもその値が 0 に近いほどデータ同化の解析精度が高い事を示す。

実験設定は,PO-EnKF・Serial EnSRF・PF でそれぞれアンサンブル数を 16 と 1000 とした. 膨張係数と局所化スケールは, 上述の実験で作成した図 1〜図 4 において、最も解析精度が高かった場合の条件を採用している. なお、詳細については考察の章で後述するが、PO-EnKF はアンサンブル数を 1000 にすると局所化を行う必要がなくなるため、これ以降の実験では膨張係数を 0.01 に設定し, 局所化は行っていない. また、本研究で用いる PF は局所化の概念を導入した LPF であるため, アンサンブル数 16 では局所化スケールを 1 に設定したが, アンサンブル数 1000 では局所化スケールを 1,10、および局所化なしに変化させ、かつ膨張係数の値も 0.75 に設定して数値実験を行った.

なお、解析 SPREAD は以下の数式に基づいて求めている。

$$SPREAD_{analysis} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\mathbf{x}^{a(i)} - \bar{\mathbf{x}}^{a})^2} \tag{3.2}$$

解析 SPREAD は、解析アンサンブル平均と解析アンサンブルの各メンバーの標準偏差である。

### 3.2.1 結果

図 5 は, データ同化手法に PO-EnKF, アンサンブル数を 16, 膨張係数を 0.05, 局所化スケールを 4 にした時の解析 RMSE と解析 SPREAD の時系列図である. 解析 RMSE の期間平均値は 0.238 で、およそ 0.1〜0.3 の範囲内で変動している。また、解析 SPREAD の期間平均値は 0.204 で、およそ 0.2 程度で変動している。それぞれの期間平均値の差が 0.1 以下である事から, 解析アンサンブルが真値を高精度で捉え続けている事が分かる.

図 6 は, データ同化手法に PO-EnKF, アンサンブル数を 1000, 膨張係数を 0.01, 局所化を行わない場合の解析 RMSE と解析 SPREAD の時系列図である. 解析 RMSE の期間平均値は 0.175 で、およそ 0.1〜0.3 の範囲内で変動している。また、解析 SPREAD の期間平均値は 0.184 で、およそ 0.2 程度で変動している。それぞれの期間平均値の差が 0.01 程度である事から, 解析アンサンブルが真値を非常に高精度で捉え続けている事が分かる.

図 7 は, データ同化手法に Serial EnSRF, アンサンブル数を 16, 膨張係数を 0.01, 局所化スケールを 7 にした時の解析 RMSE と解析 SPREAD の時系列図である. 解析 RMSE の期間平均値は 0.194 で、およそ 0.1〜0.3 の範囲内で変動している。また、解析 SPREAD の期間平均値は 0.183 で、およそ 0.2 程度で変動している。それぞれの期間平均値の差が 0.01 程度である事から, 解析アンサンブルが真値を非常に高精度で捉え続けている事が分かる.

図 8 は, データ同化手法に Serial EnSRF, アンサンブル数を 1000, 膨張係数を 0.01, 局所化スケールを 8 にした時の解析 RMSE と解析 SPREAD の時系列図である. 解析 RMSE の期間平均値は 0.228 で、およそ 0.1〜0.3 の範囲内で変動している。また、解析 SPREAD の期間平均値は 0.219 で、およそ 0.2 程度で変動している。それぞれの期間平均値の差が 0.01 程度である事から, 解析アンサンブルが真値を非常に高精度で捉え続けている事が分かる.

図 9 は, データ同化手法を PF, アンサンブル数を 16, 局所化スケールを 1 にした時の解析 RMSE と解析 SPREAD の時系列図である. 解析 RMSE の期間平均値は 0.398 で、およそ 0.3〜0.5 の範囲内で変動している。また、解析 SPREAD の期間平均値は 0.431 で、およそ 0.4〜0.6 の範囲内で変動している。それぞれの期間平均値の差が 0.1 以下である事から, 解析ア

ンサンブルが真値を高精度で捉え続けているが、全体的な精度は低い事が分かる.

図 10 は, データ同化手法を PF, アンサンブル数を 1000, 局所化スケールを 1 にした時の解析 RMSE と解析 SPREAD の時系列図である. 解析 RMSE の期間平均値は 0.349 で、およそ 0.3〜0.5 の範囲内で変動している。また、解析 SPREAD の期間平均値は 0.503 で、およそ 0.5 程度で変動している。それぞれの期間平均値の差が 0.1 以上である事から, 解析アンサンブルが真値をあまり高精度で捉えられておらず、また全体的な精度は低い事が分かる.

図 11 は, データ同化手法を PF, アンサンブル数を 1000, 局所化スケールを 10 にした時の解析 RMSE と解析 SPREAD の時系列図である. 解析 RMSE の期間平均値は 0.315 で、およそ 0.3〜0.5 の範囲内で変動している。また、解析 SPREAD の期間平均値は 0.426 で、およそ 0.4 程度で変動している。それぞれの期間平均値の差が 0.1 以上である事から, 解析アンサンブルが真値をあまり高精度で捉えられておらず、また全体的な精度は低い事が分かる.

図 12 は, データ同化手法を PF, アンサンブル数を 1000, 局所化を行わない場合の解析 RMSE と解析 SPREAD の時系列図である. 解析 RMSE の期間平均値は 4.984 で、およそ 4.0〜6.0 の範囲内で変動している。また、解析 SPREAD の期間平均値は 0.196 で、およそ 1.0〜3.0 の範囲内で変動しており、TIME=80 付近で値が 0 になっている。それぞれの期間平均値の差が 1.0 以上である事から, 解析アンサンブルは真値を捉えられていない。解析 RMSE の値が 1.0 を超えており、かつ解析 SPREAD の値が非常に低く途中で 0 になっている事から、アンサンブルの退化によってフィルタ発散に陥っている事が分かる。

図 13 は, データ同化手法を PF, アンサンブル数を 1000, 局所化スケールを 10、膨張係数を 0.75 にした時の解析 RMSE と解析 SPREAD の時系列図である. 解析 RMSE の期間平均値は 0.290 で、およそ 0.2〜0.4 の範囲内で変動している。また、解析 SPREAD の期間平均値は 0.278 で、およそ 0.3 程度で変動している。それぞれの期間平均値の差が 0.01 程度である事から, 解析アンサンブルが真値を非常に高精度で捉え続けている事が分かる.

## 3.3　非ガウス性と解析精度

解析 RMSE・解析 SPREAD・予報 KLD のホフメラー図を作成した. なお、KLD とはカルバックライブラー情報量 (KLD: Kullback–Leibler divergence) の事であり、2 つの分布の形状の違いを示す指標である。この実験では、ホフメラー図を作成する事で, 各格子点におけるデータ同化の解析精度がどの様に時間変化しているか確認した. また, 予報誤差分布とその平均と分散から得られるガウス分布の形状の違いを予報 KLD を用いて比較する事で, 各格子点におけるデータ同化の解析精度が Lorenz-96 モデルの非線形性によってどの様に変化しているか確認した. なお、解析 RMSE・解析 SPREAD・予報 KLD に単位は伴わない。それぞれの格子点の色が紫色に近いほど値が小さいためデータ同化の解析精度が高く、黄色に近いほど値が大きいためデータ同化の解析精度が低い事を示している。ここでは予報 KLD の値が高いほど,Lorenz-96 モデルの非線形性によって予報誤差分布の非ガウス性が強まり、それに伴ってデータ同化の解析精度が低下すると考える. 予報 KLD は,Lorenz-96 モデルの解が安定した後でなければ正確な検証を行う事ができない. したがって, ここでは TIME = 0〜73 つまり最初の 1 年間をスピンアップの期間とし,TIME73〜146 つまり最後の 1 年間のホフメラー図のみ作成し、その比較を行った. 実験設定は、解析精度の時間変化に関する実験を行った時と同様である。

なお、予報 KLD は以下の数式に基づいて求めている。

$$KLD_{forecast} = \int p(x) \log \frac{p(x)}{q(x)} dx \tag{3.3}$$

KLD は、2 つの確率密度分布の形状がどれほど異なっているかを示す指標である。

$$p(x) = p(\mathbf{x}_t \mid \mathbf{y}_{1:t-1}) \tag{3.4}$$

$$q(x) = \frac{1}{\sqrt{2\pi\sigma}} exp\left(-\frac{(\mathbf{x}^f - \mu)^2}{2\sigma^2}\right) \tag{3.5}$$

ここで $p(x)$ は予報アンサンブルの分布であり、$q(x)$ は予報アンサンブルの平均と分散から作成したガウス分布である。

### 3.3.1 結果

図 14 は, データ同化手法に PO-EnKF, アンサンブル数を 16, 膨張係数を 0.05, 局所化スケール 4 の時の解析 RMSE・解析 SPREAD・予報 KLD のホフメラー図である. 解析 RMSE と解析 SPREAD の分布には似た傾向が見られるのに対し、予報 KLD の分布からはあまり特徴を見出す事ができない。ここから、この条件でデータ同化を行った場合、解析 RMSE の変動は解析 SPREAD に依存しやすいと考えられる。

図 15 は, データ同化手法に PO-EnKF, アンサンブル数を 1000, 膨張係数を 0.01, 局所化を行わない場合の解析 RMSE・解析 SPREAD・予報 KLD のホフメラー図である. 解析 RMSE と解析 SPREAD の分布には似た傾向が見られるのに対し、予報 KLD の分布はほぼ一定して紫色であり、非ガウス性が低い事を示している。ここから、この条件でデータ同化を行った場合、解析 RMSE の変動は解析 SPREAD にかなり依存すると考えられる。

図 16 は, データ同化手法に Serial EnSRF, アンサンブル数を 16, 膨張係数を 0.01, 局所化スケールを 7 にした時の解析 RMSE・解析 SPREAD・予報 KLD のホフメラー図である. 解析 RMSE と解析 SPREAD と比較すれば解析 KLD の分布から似た傾向を見出す事は難しいが、全体的にそれぞれの分布が時間経過と共に左上から右下へ推移している。ここから、この条件でデータ同化を行った場合、解析 RMSE の変動は解析 SPREAD と解析 KLD に依存しやすく、予報誤差分布の非ガウス性が強まるとデータ同化の解析精度が低下すると考えられる。

図 17 は, データ同化手法に Serial EnSRF, アンサンブル数を 1000, 膨張係数を 0.01, 局所化スケールを 8 にした時の解析 RMSE・解析 SPREAD・予報 KLD のホフメラー図である. 解析 RMSE と解析 SPREAD の分布には似た傾向が見られるのに対し、予報 KLD の分布は常に黄色であり、非ガウス性が高い事を示している。ここから、この条件でデータ同化を行った場合、解析 RMSE の変動は解析 SPREAD にかなり依存し、かつ予報 KLD の低下が解析 RMSE に悪影響を及ぼすと考えられる。

図 18 は, データ同化手法に PF, アンサンブル数を 16, 局所化スケール 1 の時の解析 RMSE・解析 SPREAD・予報 KLD のホフメラー図である. 解析 RMSE と解析 SPREAD の分布には

似た傾向が見られるのに対し、予報 KLD の分布からはあまり特徴を見出す事ができない。ここから、この条件でデータ同化を行った場合、解析 RMSE の変動は解析 SPREAD に依存しやすいと考えられる。

図 19 は, データ同化手法に PF, アンサンブル数を 1000, 局所化スケール 1 の時の解析 RMSE・解析 SPREAD・予報 KLD のホフメラー図である. 解析 RMSE の分布は時間変化しているのに対し、解析 SPREAD の分布はほぼ一定して黄色、解析 KLD の分布はほぼ一定して紫色であり、解析アンサンブルの広がりは大きいが、非ガウス性は低い事を示している。ここから、この条件でデータ同化を行った場合、解析 RMSE の変動は解析 SPREAD と解析 KLD に依存しやすいと考えられる。

図 20 は, データ同化手法に PF, アンサンブル数を 1000, 局所化スケール 10 の時の解析 RMSE・解析 SPREAD・予報 KLD のホフメラー図である. 解析 RMSE と解析 SPREAD の分布には似た傾向が見られるのに対し、予報 KLD はほぼ一定して紫色であり、非ガウス性は低い事を示している。ここから、この条件でデータ同化を行った場合、解析 RMSE の変動は解析 SPREAD にかなり依存すると考えられる。

図 22 は, データ同化手法に PF, アンサンブル数を 1000, 局所化スケール 10、膨張係数 0.75 の時の解析 RMSE・解析 SPREAD・予報 KLD のホフメラー図である. 解析 RMSE と解析 SPREAD の分布には似た傾向が見られるのに対し、予報 KLD はほぼ一定して紫色であり、非ガウス性は低い事を示している。ここから、この条件でデータ同化を行った場合、解析 RMSE の変動は解析 SPREAD にかなり依存すると考えられる。

## 3.4 観測密度と解析精度

PF は EnKF と違ってガウス分布の仮定をしないため、非線形性の強い現象に対しては EnKF より有効に機能する、という考えに基づいて研究を行ってきたが、上述した結果は全て PF より EnKF の方がデータ同化の解析精度が高いという事を示していた。しかし、予報誤差分布の非ガウス性が高い場合であれば、PF の方が解析精度が高くなる可能性がある。そ

こで、観測点の疎密を作成し、人為的に非ガウス性が高くなる状況を再現した。ここで、観測点が疎な領域は現実の海洋、観測点が密な領域は現実の大陸に相当すると考える。実験設定としては、格子点番号 1〜20 には観測点が存在するが、格子点番号 21〜40 の間には観測点が格子点番号 31 にのみ存在する。この様な状況で、解析 RMSE・解析 SPREAD の時系列図と、解析 RMSE・解析 SPREAD・予報 KLD のホフメラー図を作成し、PF の解析精度が EnKF より高くなるか確認する。なお、観測密度以外の実験設定は上述したものと同様である。

### 3.4.1 結果

図 23 は, 観測点の疎密を与えた状態で、データ同化手法に PO-EnKF, アンサンブル数を 1000, 膨張係数を 0.01, 局所化を行わない場合の解析 RMSE と解析 SPREAD の時系列図である. 解析 RMSE の期間平均値は 1.580 で、およそ 1.0〜3.0 の範囲内で変動している。また、解析 SPREAD の期間平均値は 1.901 で、およそ 2.0 程度で変動している。観測誤差の RMSE は 1.0 で、解析 RMSE がそれより高い値になっているため、フィルタ発散に陥っている。しかし、それぞれの期間平均値の差が 1.0 以下であり、データ同化システムは機能し続けている事から, 解析アンサンブルが真値を低い精度で捉え続けている事が分かる.

図 24 は, 観測点の疎密を与えた状態で、データ同化手法を PF, アンサンブル数を 1000, 局所化スケールを 10、膨張係数を 0.75 にした時の解析 RMSE と解析 SPREAD の時系列図である. 解析 RMSE の期間平均値は 2.143 で、およそ 1.0〜3.0 の範囲内で変動している。また、解析 SPREAD の期間平均値は 1.597 で、およそ 2.0 程度で変動している。観測誤差の RMSE は 1.0 で、解析 RMSE がそれより高い値になっているため、フィルタ発散に陥っている。しかし、それぞれの期間平均値の差が 1.0 以下であり、データ同化システムは機能し続けている事から, 解析アンサンブルが真値を低い精度で捉え続けている事が分かる.

図 25 は、PO-EnKF と PF の解析精度を比較しやすくする為に、それぞれの解析 RMSE のみ描画した時系列図である。全体的に PO-EnKF の解析 RMSE の方が PF よりも低く、解析精度が高い。しかし、時刻によっては PF の解析 RMSE の方が PO-EnKF よりも低い。

図 26 は, 観測点の疎密を与えた状態で、データ同化手法に PO-EnKF, アンサンブル数を 1000, 膨張係数を 0.01, 局所化を行わない場合の解析 RMSE・解析 SPREAD・予報 KLD のホフメラー図である. 観測点が疎な図の右側ではそれぞれの図が黄色になっており、解析 SPREAD と予報 KLD の値が高くなる事で、解析精度が低くなっている事が分かる。

図 27 は, 観測点の疎密を与えた状態で、データ同化手法に PF, アンサンブル数を 1000, 局所化スケール 10、膨張係数 0.75 の時の解析 RMSE・解析 SPREAD・予報 KLD のホフメラー図である. 観測点が疎な図の右側では解析 RMSE と解析 SPREAD の図が黄色になっているが、予報 KLD の図は右側でも紫色になっている。ここから、主に解析 SPREAD の値が高くなる事で、解析精度が低くなっている事が分かる。

# 第4章　考察

## 4.1　最適な解析条件の推定に関する考察

図 1 と図 2 を比較すると、アンサンブル数の増加によって膨張係数と局所化スケールの比例関係を表した境界線が解消され、フィルタ発散に陥る事が無くなっている。これは、アンサンブル数が小さいとサンプリングエラーによってデータ同化の解析精度が低下する為である。

また、図 3 と図 4 を比較すると、PO-EnKF の時ほど図は大きく変化しない。PO-EnKF はアンサンブル数が増加すると解析精度が向上するが、Serial EnSRF はアンサンブル数が増加すると逆に解析精度が低下する。これはアンサンブルアップデートの違いが影響しており、PO-EnKF では観測に摂動を加えて解析アンサンブルを作成するが、Serial EnSRF は解析誤差共分散行列から直接解析アンサンブルを作成する。したがって、Serial EnSRF ではアンサンブルの外れ値が生じると元へ戻り難く、予報誤差分布の非ガウス性が強まる。この傾向はアンサンブル数が増加するほど顕著になるため、アンサンブル数と解析精度は必ずしも比例しない (Anderson 2010, Amezcua et al. 2012) 。ただし、Serial EnSRF は小さいアンサンブル数でも解析精度が高いため、計算機資源が限られている事を考慮すれば、コストパフォーマンスが高い手法である。

更に、PO-EnKF と Serial EnSRF に共通した傾向として、アンサンブル数が増加すると、膨張係数が小さく局所化スケールが大きいほど解析精度が向上する。これは、膨張係数を大きくするとアンサンブルが真値を捉えやすくなるが、その分アンサンブルの広がりが大きくなり、解析精度が低下する為である。また、アンサンブル数が小さい時は局所化スケールを

大きくしてサンプリングエラーを抑える必要があるが、アンサンブル数が大きい時に同じ様にすると、本来扱うべき誤差共分散も無視されてしまう。したがって、アンサンブル数が大きい時は、局所化スケールも大きくした方がデータ同化の解析精度は向上しやすいと考えられる。。

## 4.2 解析精度の時間変化に関する考察

図 14 と図 15 を比較すると、アンサンブル数の増加によって解析 RMSE と解析 SPREAD の変動幅が小さくなり、解析精度が向上している事が分かる。また、図 7 と図 8 を比較すると、アンサンブル数を増加させてにも関わらず、解析精度が向上しておらず、むしろ僅かに解析精度が低下している事が分かる。ここから、PO-EnKF ではアンサンブル数と解析精度が比例するが、Serial EnSRF では比例しない事が分かる。これは、上述した様にアンサンブルアップデートの方法が異なる為である。しかし、アンサンブル数をおよそ 100 倍にしても解析 RMSE が 0.1 程度しか減少しないため、PO-EnKF は効率的とは言えない。対して、Serial EnSRF はアンサンブル数が小さくても解析精度が高く現業予報に向いており、実際に気象庁で Serial EnSRF の並列計算効率を改善した LETKF が採用されている事を考えれば、この結果は妥当である。

図 9 と図 10 を比較すると、粒子数の増加によって解析 RMSE と解析 SPREAD の変動幅がかなり小さくなり、かつ解析精度が向上している事が分かる。特に、解析 SPREAD に注目するとその傾向が顕著であり、PF は EnKF に比べて粒子数の影響が強く現れる事が分かる。

次に、図 10 と図 11 を比較すると、局所化スケールの増加によって解析精度が向上しており、かつ解析 SPREAD の値が解析 RMSE に近くなっている。これは、上述した様にアンサンブル数が大きい時ほど局所化スケールを大きくした方が、誤差共分散を正確に求められる様になる為であり、この傾向は EnKF だけでなく PF にも共通する事が分かる。ただし、図 12 では局所化を行わない様にすると、アンサンブルの退化によってフィルタ発散に陥る事が示されている。これは、局所化を行うと各格子点で同化される観測が異なるため粒子の

多様性が維持されるが、局所化を行わないと全格子点で同化される観測が等しくなる事で粒子全体が似た構造になり、アンサンブルの退化が生じやすくなる為だと考えられる。この様に、PF は局所化に対して非常に敏感であり、ただ粒子数を大きくするだけでなく、局所化スケールを適切な値にする事がかなり重要となる。ただし、現在でも局所化の動的推定は実現されていないため、もしも PF を現業予報で用いるとなれば、これは解決すべき 1 つの問題になるであろう。また、フィルタ発散に陥った図 12 を除けば、PF はいずれの場合でも解析 SPREAD の値が解析 RMSE を上回っており、これはアンサンブルの広がりが過剰という事を示している。これは、おそらく PF の特徴であり、解析 SPREAD が過剰に大きいと解析精度は低下するため、PF の性能を向上させるためには、共分散膨張の概念を導入し、膨張係数を適切な値にする必要がある。

そこで、図 13 では共分散膨張に相当する概念を PF に導入し、膨張係数の値をチューニングした。すると、解析 RMSE と解析 SPREAD の変動幅が更に小さくなり、かつ解析精度が向上した。更に、解析 SPREAD の値が解析 RMSE とほぼ等しくなり、アンサンブルが真値をより正確に捉える様になった。上述の実験では、全て膨張係数の値が 0.5 に設定されているが、膨張係数を 0.75 にした事で特定の粒子の重みが相対的に大きくなった。これにより、粒子全体が真値に近い値を取る様になり、解析精度が向上した。上述した様に、局所化スケールを動的推定できる様にする事が PF の性能を向上させる為の 1 つのポイントだが、これは膨張係数にも言える事である。膨張係数の動的推定は EnKF では既に実現されているため、PF でも同様の事が実現できれば、性能は大きく向上するであろう。しかし、それでも図 15 と図 13 を比較すると PF の解析精度の方が低く、PF が EnKF と同等以上の性能になる為には、上述した以外の問題も解決する必要がある。

## 4.3 非ガウス性と解析精度に関する考察

図 14 と図 15 を比較するとアンサンブル数の増加によって予報 KLD の図がより紫色になっており、予報誤差分布の非ガウス性が低くなっている事が分かる。これは、単純にアンサンブル数が増加した事で、予報誤差分布がよりガウス分布に近い滑らかな形状になった為だと考えられる。

図 16 と図 17 を比較するとアンサンブル数の増加によって予報 KLD の図が黄色になっており、予報誤差分布の非ガウス性がかなり強まっている事が分かる。これは、上述した様に Serial EnSRF ではアンサンブルの外れ値が生じやすいため、アンサンブル数を大きくした事でむしろ外れ値が増加し、非ガウス性が強まったと考えられる。また、図 14 と図 16 を比較すると、Serial EnSRF における解析 RMSE・解析 SPREAD・予報 KLD の図の方がより全体的に一致している。これは、解析 RMSE の値が解析 SPREAD と予報 KLD の影響を受ける為であり、実際に Kondo and Miyoshi (2019) では予報 KLD の値が高くなると解析 RMSE の値も高く事が明らかになっている。また、現実大気では熱帯の積雲対流や低気圧の周辺で, 大気の非線形性によって予報誤差分布の非ガウス性が強まり、それによって気象予測の精度が低くなる事が知られている (Kondo and Miyoshi 2019, Kawabata and Ueno 2020)。これは、予報誤差が非ガウス分布になると EnKF におけるガウス分布の仮定が満たされなくなる事に起因しており、PO-EnKF と違って観測に摂動を加えない Serial EnSRF ではアンサンブルの外れ値が発生しやすいため、非ガウス性に弱いという EnKF の特徴がより明らかになったと考えられる。

図 18 と図 19 を比較すると、粒子数の増加によって PO-EnKF と同様に予報 KLD の図がより紫色になっているが、解析 SPREAD の図がより黄色になっている。これは、予報誤差分布の非ガウス性は低くなったが、アンサンブルの広がりが大きくなった事を示しており、粒子数の増加によって粒子の多様性も増加した為だと考えられる。PF は、ガウス分布の仮定をしないため、誤差分布を作成するにあたって大量の粒子を必要とする。したがって、粒子数が充分大きくないと機能せず、解析精度を高くする為には粒子数を大きくする事が必要となる。しかし、ただ単純に粒子数を大きくすれば良い訳ではなく、解析精度を更に向上させる為には局所化や膨張係数といった他のチューニングパラメータも適切な値に設定する必要がある。

図 19 と図 20 を比較すると、局所化スケールの増加によって解析 SPREAD の図はより紫色になっているが、予報 KLD の図はより黄色になっている。これは、局所化スケールの増加によって観測の情報がより多く同化され、粒子全体が真値をより捉える様になった分、特定の粒子に偏る様になり、予報誤差の非ガウス性が高くなった為だと考えられる。

図 21 においてアンサンブルの退化によりフィルタが発散した事もこれに起因していると考えられ、PF は PO-EnKF と違って局所化スケールをただ単純に大きくすれば良い訳ではない。PO-EnKF では解析精度がアンサンブル数と局所化スケールに比例したにも関わらず、PF で同じ様にならないのはデータ同化の方法が異なる為だと考えられる。EnKF では、アンサンブル全体を観測に近づけるため、アンサンブル数を無限大にすると KF と同値になるため、解析精度が向上する。特に、PO-EnKF はアンサンブル数を大きくするほど観測に加えた摂動に起因するサンプリングエラーが減少するため、この傾向が顕著になる。一方、PF では観測に近い粒子を複製し、観測から遠い粒子を破棄するため、単純にアンサンブル数と局所化スケールを大きくしても、観測から遠い粒子が増えるだけや、特定の粒子に偏りが生じるだけで、解析精度自体はあまり向上しないという事が起こりうると考えられる。

それらを踏まえた上で図 22 と図 20 を比較すると、解析 SPREAD の図がより紫色になっている。これは、膨張係数を大きくして粒子の多様性を抑えた事で、粒子全体が真値をより効率的に捉える様になり、アンサンブルの広がりの低下と共に解析精度が向上したと考えられる。上述してきた事をまとめると、EnKF は PF と違ってアンサンブル数などの増加と解析精度が単純な比例関係になく、アンサンブル数・局所化スケール・膨張係数といった、様々なチューニングパラメータのバランスによって解析精度が決まるという事になる。これは、PF がアンサンブルの退化という問題を抱えているため、粒子全体が真値を捉え、かつ粒子の多様性が失われない様に、慎重にチューニングパラメータを設定する必要があるという事になる。

## 4.4　観測密度と解析精度に関する考察

図 23 と図 24 を比較すると、解析 SPREAD の変動幅はほぼ等しいが、解析 RMSE の変動幅は PF の方が PO-EnKF よりも大きく、データ同化システムの挙動が不安定である。解析 RMSE は PF の方が PO-EnKF よりも全体的に高いのに対し、解析 SPREAD は PF の方が PO-EnKF よりも全体的に低く、解析アンサンブルの広がりが過剰に小さいため、解析精度が低くなっていると考えられる。

また、図 25 を見ると PF の方が PO-EnKF よりも解析 RMSE が全体的に高く、解析精度が低くなっているが、時刻によっては同等以上になっている時もあるため、チューニングパラメータの設定次第では PF の方が PO-EnKF よりも解析精度が高くなる可能性がある。これは、観測点の疎密を与える前の図 6 と図 13 には見られなかった特徴である。

図 26 と図 27 を比較すると、PF の方が PO-EnKF よりも解析 RMSE と解析 SPREAD の図が黄色になっている。ホフメラー図では全体的に解析 SPREAD が大きいにも関わらず、時系列図では小さくなっているという事は、一部の格子点においてアンサンブルの退化が生じている可能性がある。したがって、PF の解析精度を向上させる為には、膨張係数を小さくして粒子の多様性を増加させ、粒子全体が真値を捉えやすくすると良い可能性がある。また、注目すべき点として PF では観測点が少ないにも関わらず、予報 KLD の図の右側が紫色になっており、PO-EnKF よりも非ガウス性が小さくなっている。ここから、PF はガウス分布の仮定をしないため、非ガウス性の影響を抑える働きがあり、大陸に比べて観測点が少ない海洋などでは、PO-EnKF よりも有効になると考えられる。しかし、その時の膨張係数は小さくして、観測が少ない状態でも粒子全体が真値を捉えやすくする必要があると考えられる。

# 第5章　結論

PO-EnKF は、アンサンブル数を大きくすると解析精度が向上するのに対し、Serial EnSRF はアンサンブル数を大きくすると解析精度が低下する事が分かった。これは、アンサンブルアップデートの方法の違いに起因しており、アンサンブル数が増加すると、PO-EnKF ではサンプリングエラーが減少するが、Serial EnSRF ではアンサンブルの外れ値が生じる為である。また、アンサンブル数が大きい場合は、どちらも局所化スケールを大きくした方が解析精度が高くなる。しかし、アンサンブル数が小さい場合は Serial EnSRF の方が PO-EnKF よりも解析精度が高く、コストパフォーマンスが高いため現業予報に向いている。ただし、非ガウス性によって解析精度が低下しやすい。

PO-EnKF と PF は、アンサンブル数が大きい方が解析精度が高くなるが、PF はその特徴がより顕著である。これは、アンサンブル数が大きいと誤差分布の形状が滑らかになり、粒子全体が真値を捉えやすくなる為である。ただし、単純にアンサンブル数を増加させるだけでは、解析精度が向上し難い。また、PF ではアンサンブルの退化が生じるため、解析精度がチューニングパラメータに対して敏感であり、粒子全体が真値を捉えられる様にしつつ、かつアンサンブルの退化が生じない様にしなければならない。そのため、局所化スケールや膨張係数の動的推定を実現しなければ、現業予報に用いる事はできない。

EnKF ではガウス分布の仮定をしているため、大気の非線形性が強い顕著現象の下ではそれが満たされず、解析精度が低下する。しかし、PF はその様な仮定をしないため、その様な状況でも有効に機能する可能性がある。この様な考えに基づいて、Lorenz-96 モデルにおける EnKF と PF の解析精度を比較したが、PF の解析精度が EnKF を上回る事は無かった。そこで、観測点の疎密を与えて非ガウス性を強くしたところ、PF と PO-EnKF の解析精度が近い値になり、チューニングパラメータの設定次第では、PF の解析精度の方が高くなる可能性が示唆された。PF はガウス分布の仮定をしないため、観測点が少ない海洋などでも

解析精度が低下し難く、EnKF よりも有効なデータ同化手法になりうる。

# 謝辞

本研究を進めるにあたって, 指導教員である筑波大学計算科学研究センター (CCS) の田中博教授には, 研究手法や結果に対する考察などについて終始適切な御指導を賜り, 心から感謝しております. 特に, 本研究室の先輩にもあたる気象庁気象研究所の近藤圭一博士から直接ご指導を賜る機会をご用意して頂いた事には, 誠に感謝しております. また Lorenz-96 モデルの構造についてご質問した際には, 田中先生の深い見識に基づく貴重なご意見を述べて頂き, どうもありがとうございます. 気象庁気象研究所の近藤圭一博士には, 研究会や打ち合わせを通じて, データ同化に関する専門的知識を持ち合わせた, 親切かつ丁寧な御指導を全面的に賜ったこと深く感謝致します. 近藤圭一博士の御指導なくして, 私がこの研究を進める事はできませんでした. 筑波大学生命環境系の日下博幸教授, 植田宏昭教授, 上野健一准教授, 松枝未遠助教, 釜江陽一助教, 原田真理子助教, ドアングアンヴァン助教, 中村祐輔特任助教には, 普段より研究を進める為の礎となる学問の指導に加え, 発表の機会を通じて非常に参考になるコメントをして頂きましたこと, 誠に感謝致します. また, 本研究室の大学院生の豊岡大地さん, 石山涼太さん, 縄司瑛太さん, 熊谷怜皇さん, 伊藤駿さんには, 研究を進めるにあたり様々な質問や相談に乗って頂き, 時には貴重なアドバイスや議論をして頂いたこと, 心から感謝致します.. 今では田中先生の研究室に配属された事を, 本当に嬉しく思っています. 最後に, 共に研究を進めた筑波大学生命環境科学研究科大気分野の研究員, 博士課程, 修士課程の先輩方, 筑波大学生命環境学群地球学類大気科学分野の共に学士論文執筆作業を進めた 4 年生の皆様には, 時に協力し合い, 励まし合い, 高め合えたこと, 心より感謝致します.

本論文は以上の皆様のご協力により完成させる事ができました, 心より深く感謝致します.

# 参考文献


近藤 圭一, 2009: NICAM-LETKF の開発および 4 次元データ同化実験, 筑波大学大学院 生命環境科学研究科 地球科学専攻 修士論文.

露木 義, 1997: 変分法によるデータ同化. 数値予報課報告・別冊 第 43 号, 気象庁予報部, 103-165.

露木 義, 川畑 拓矢, 2008: 気象学におけるデータ同化. 気象研究ノート 第 217 号, 日本気象学会.

西村 和裕, 2011: アンサンブルカルマンフィルタを用いた簡易大気モデルでの感度解析法の比較, 京都大学大学院 理学研究科 地球惑星科専攻 修士論文.

樋口 知之, 上野 玄太, 中野 慎也, 中村 和幸, 吉田 亮, 2011: シリーズ ¡予測と発見の科学¿ 6, データ同化入門-次世代のシミュレーション技術-, 朝倉書店.

三好 建正, 2006: アンサンブル技術の短期・中期予報への利用 〜激しい気象現象の予測向上を目指して〜, 数値予報課報告・別冊, 第 52 号, 80-99.

村上 茂教, 1997: 上部成層圏解析, 数値予報課報告・別冊, 第 43 号, 87-101.

Amezcua, J., K. Ide., C. H. Bishop., E. Kalney, 2012: Ensemble clustering in deterministic ensemble Kalman filters. *Tellus*, **64**, 1-12. doi: 10.3402/tellusa.v64i0.18039.

Anderson, J. L., 2010: A Non-Gaussian Ensemble Filter Update for Data Assimilation. *Mon. Wea. Rev.*, **138**, 4186–4198. doi: 10.1175/2010MWR3253.1.

Andrew, A., 1986: A square root formulation of the Kalman covariance equations. *AIAA J.*, **6**, 1165-1168.

Charney, J., M. Halem, and R. Jastrow, 1969: Use of incomplete historical data to infer the present state of the atmosphere. *J. Atmos. Sci.*, **26**, 1160-1163.

Corazza, M., Kalnay, E., Patil, D., J, Yang S-C., Morss, R., CaiM, SzunyoghI., Hunt, B. R., Yorke, J. A., 2003: Use of the breeding technique to estimate the structure of the analysis "error of the day". *Nonliner Proc. Geophys.*, **10**, 233-243.

Dee, D. P., 1995: On-line estimation of error covariance parameters for Atmosheric data assmilation. *Mon. Wea. Rev.*, **123**, 1128–1145.

Eliassen, A. 1954: *Provisional report on calculation of spatial covariance and autocorrelation of the pressure field. No. 5*, Videnskaps-Akademiets Institutt for Vaer-Og Klimaforsking, Oslo. (reprinted in Bengtsson et al. 1981, pp. 319–330)

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

Gandin, L. S. 1963: *Objective analysis of meteorological fields*, Gidrometeorologicheskoe Izdatelstvo, Leningrad (in Russian), English traslation by Israeli Program for Scientific Translation, Jerusalem, 1-965.

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

Gilchrist, B. and G. Cressman 1954: An experiment in objective analysis. *Tellus*, **6**, 309–318.

Hoke, J. and R. Anthes, 1976: The initialization of numerical model by a dynamic initialization technique. *Mon. Wea. Rev.*, **104**, 1551-1556.

Honda, T., S. Takino., and T. Miyoshi, 2019: Improving a Precipitation Forecast by Assimilating All-Sky Himawari-8 Satellite Radiances: A Case of Typhoon Malakas (2016). *SOLA*, Vol. **15**, 7-11, doi:10.2151/sola.2019-002.

Houtekamer, P. L., and H. L. Mitchell, 1998: data assimilation using an ensemble Kalman filter technique. *Mon. Wea. Rev.*, **126**, 796–811.

Hunt, B. R., 2005: Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter. arXiv: physics/0511236v1, 25pp.

Jazwinski, A. H., 1970: *Stochastic Processes and Filtering Theory.* Academic Press, 376pp.

Kalman, R. E., 1960: A new approach to linear filtering and predicted problems. *Trans. ASME*, *Ser. D*, *J. Basic Eng.*, **82**, 35–45.

Kawabata, T. and G. Ueno, 2020: Non-Gaussian Probability Densities of Convection Initiation and Development Investigated Using a Particle Filter with a Storm-Scale Numerical Weather Prediction Model. *Mon. Wea. Rev.*, **148**, 3-20. doi: 10.1175/MWR-D-18-0367.1.

Kitagawa, G., 1996: Monte Carlo filter and smoother for non-Gaussian nonlinear state space models. *J. Comput. Graph. Stat.*, **5**, 1–25.

Kondo, K. and T. Miyoshi, 2016: Impact of removing covariance localization in an ensemble Kalman filter: experiments with 10,240 members using an intermediate AGCM. *Mon. Wea. Rev.*, **144**, 4849-4865. doi:10.1175/MWR-D-15-0388.1.

Kondo, K., and T. Miyoshi, 2019: Non-Gaussian statistics in global atmospheric dynamics: a study with a 10240-member ensemble Kalman filter using an intermediate atmospheric general circulation model. *Nonlinear Processes in Geophys.*, **26**, 211-225. doi:10.5194/npg-26-211-2019.

Li, H., E. Kalnay and T. Miyoshi 2009: Simultaneous estimation of covariance inflation and observation errors within ensemble Kalman filter. *Quart. J. Roy. Meteor. Soc.*, **135**, 523-533. doi: 10.1002/qj.371.

Liu, H. and X. Zou, 2001: The impact of NORPEX targeted dropsondes on the analysis and 2-3-day forecasts of a landfalling Pacific winter storm using NCEP 3D-Var and 4D-Var systems. *Amer. Meteor. Soc.*, **129**, 1987–2004.

Lorenz, E. N. 1996: Predictability: A problem partly solved. ECMWF, Seminar on Predictability, 4-8 September 1995.

Miyoshi, T. and K. Aranami, 2006: Applying a Four-dimensional Kalman Filter (4DLETKF) to the JMA Nonhydrostatic Model (NHM). *SOLA*, **2**, 128-131.

Miyoshi, T., E. Kalnay, and B. Hunt, 2005: Ensemble Kalman filter expreiments with the Lorenz-96 model in preparation.

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. doi:10.1175/2010MWR3570.1

Panofsky, H. 1949: Objective wather-map analysis. *J. Appl. Meteor.*, **6**, 386–392.

Parrish, D. F. and J. C. Derber, 1992: The National Meteorological Center ' s spectral statistical-interpolation analysis system. *Mon. Wea. Rev.* **120**, 1747–1763.

Pham, D. T., J. Verron, and M. C. Roubaud, 1998: A singular evolutive extended Kalman filter for data assimilation in oceanography. *J. Mar. Syst.*, **16**, 323–340.

Thompson, P., 1961: Dynamical method of analyzing meteorological data. *Tellus*, **13**, 334-349.

van Kekem, D. L., 2018: Dynamics of the Lorenz-96 model: Bifurcations, symmetries and waves. [Groningen]: Rijksuniversiteit Groningen.

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

Whitaker, J. S. and T.M. Hamill, 2012: Evaluating Methods to Account for System Errors in Ensemble Data Assimilation. *Mon. Wea. Rev.*, **140**, 3078–3089.

Zhang, F., C, Snyder, and J. Sun, 2004: Impact of initial estimate and observation availability onconvective-scale data assimilation with an ensemble Kalman filter. *Mon. Wea. Rev.*, **132**, 1238-1253.

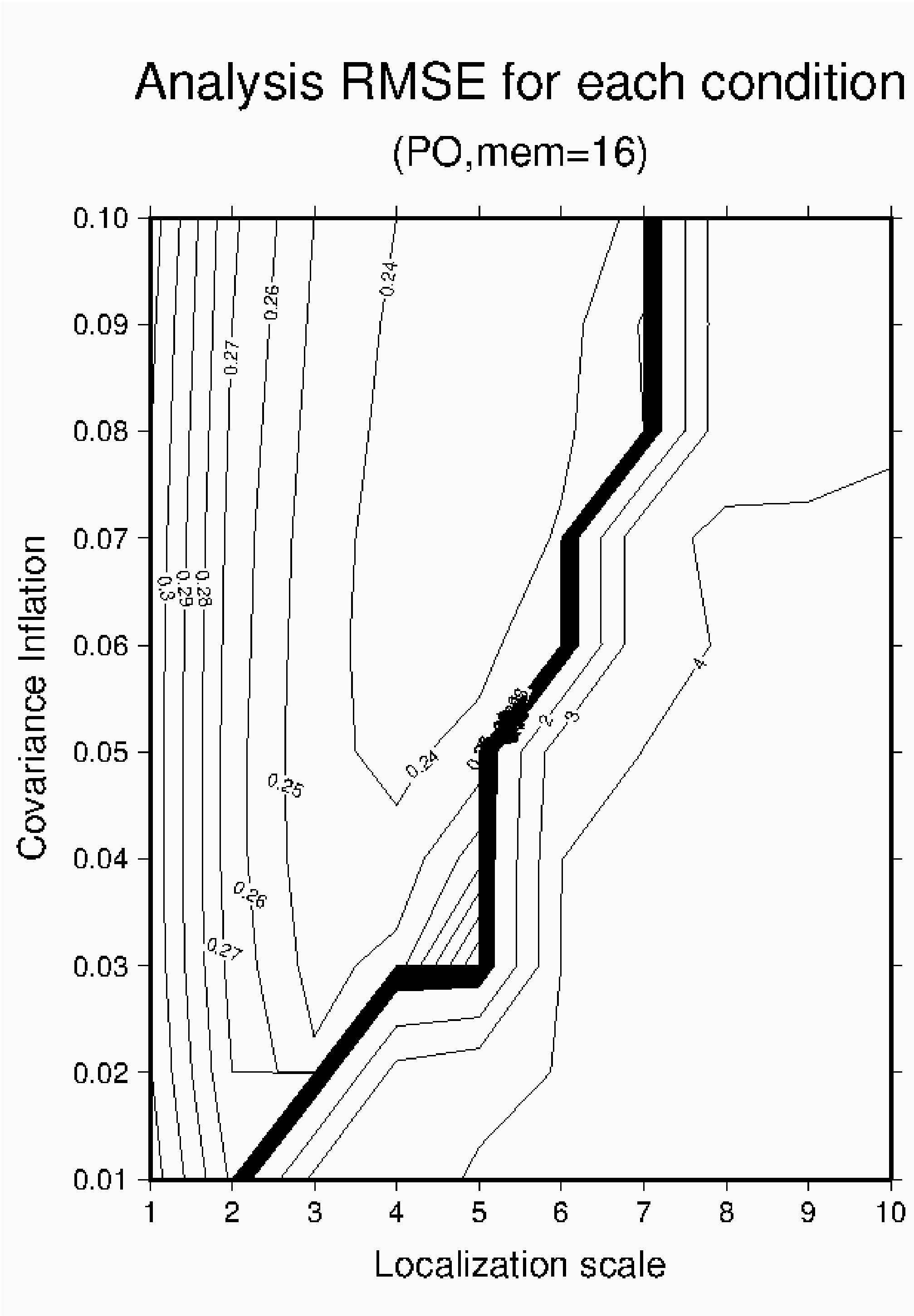


図 1: データ同化手法は PO-EnKF, アンサンブル数 16 の時の真値と解析値の解析 RMSE の等値線図. 横軸は局所化スケール, 縦軸は膨張係数.

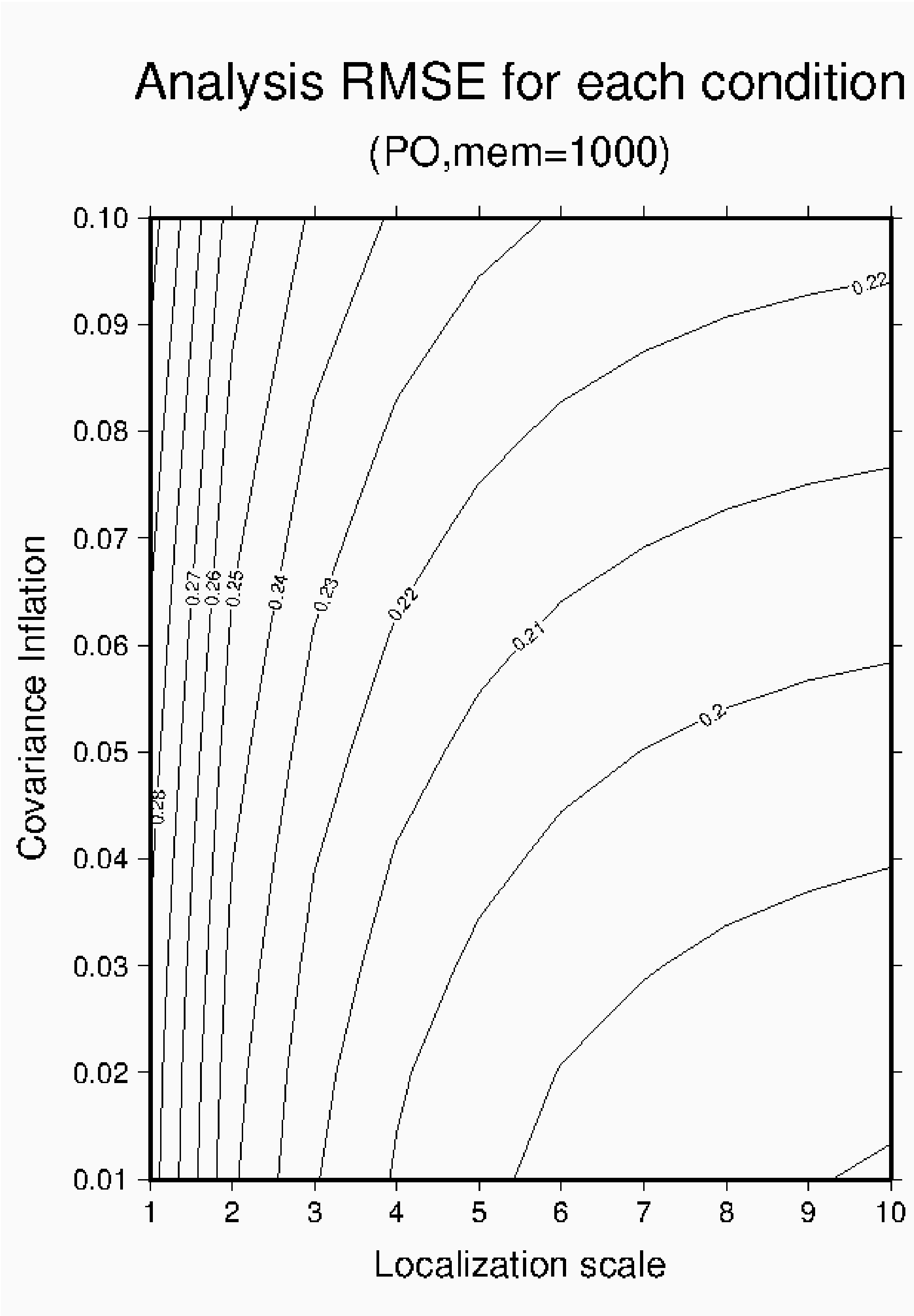


図 2: データ同化手法は PO-EnKF, アンサンブル数 1000 の時の真値と解析値の解析 RMSE の等値線図. 横軸は局所化スケール, 縦軸は膨張係数.

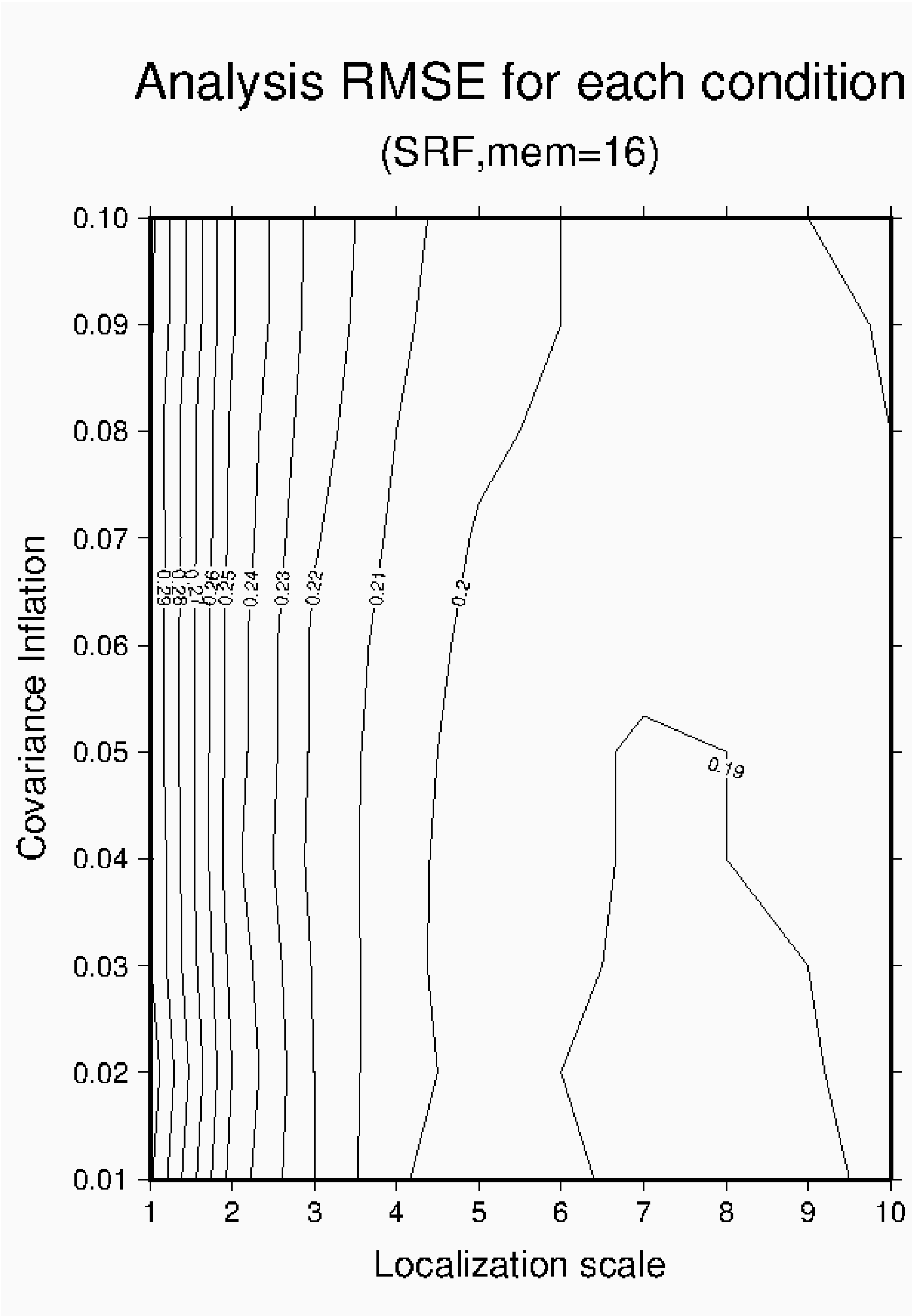


図 3: データ同化手法は Serial EnSRF, アンサンブル数 16 の時の真値と解析値の解析 RMSE の等値線図. 横軸は局所化スケール, 縦軸は膨張係数.

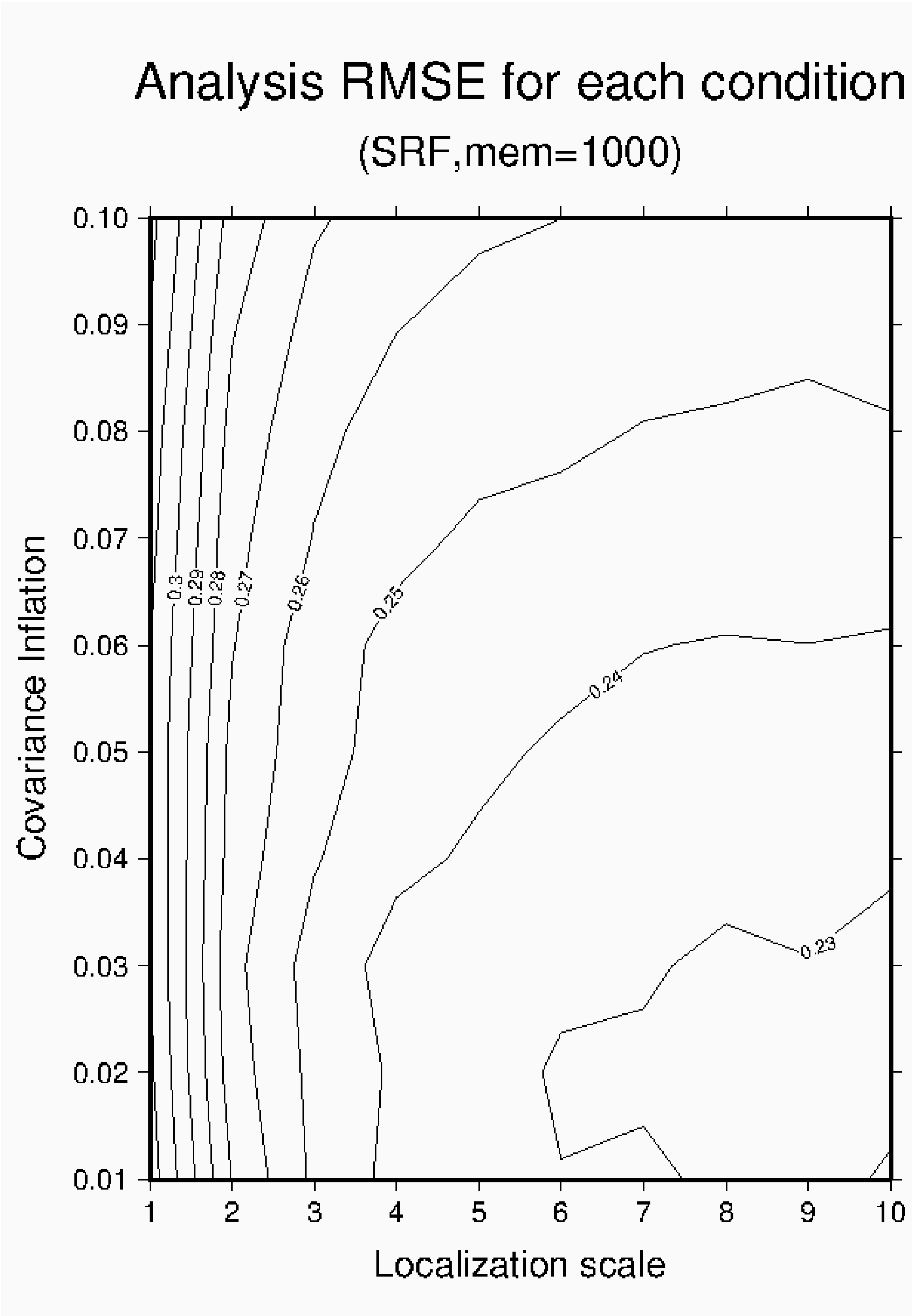


図 4: データ同化手法は Serial EnSRF, アンサンブル数 1000 の時の真値と解析値の解析 RMSE の等値線図. 横軸は局所化スケール, 縦軸は膨張係数.

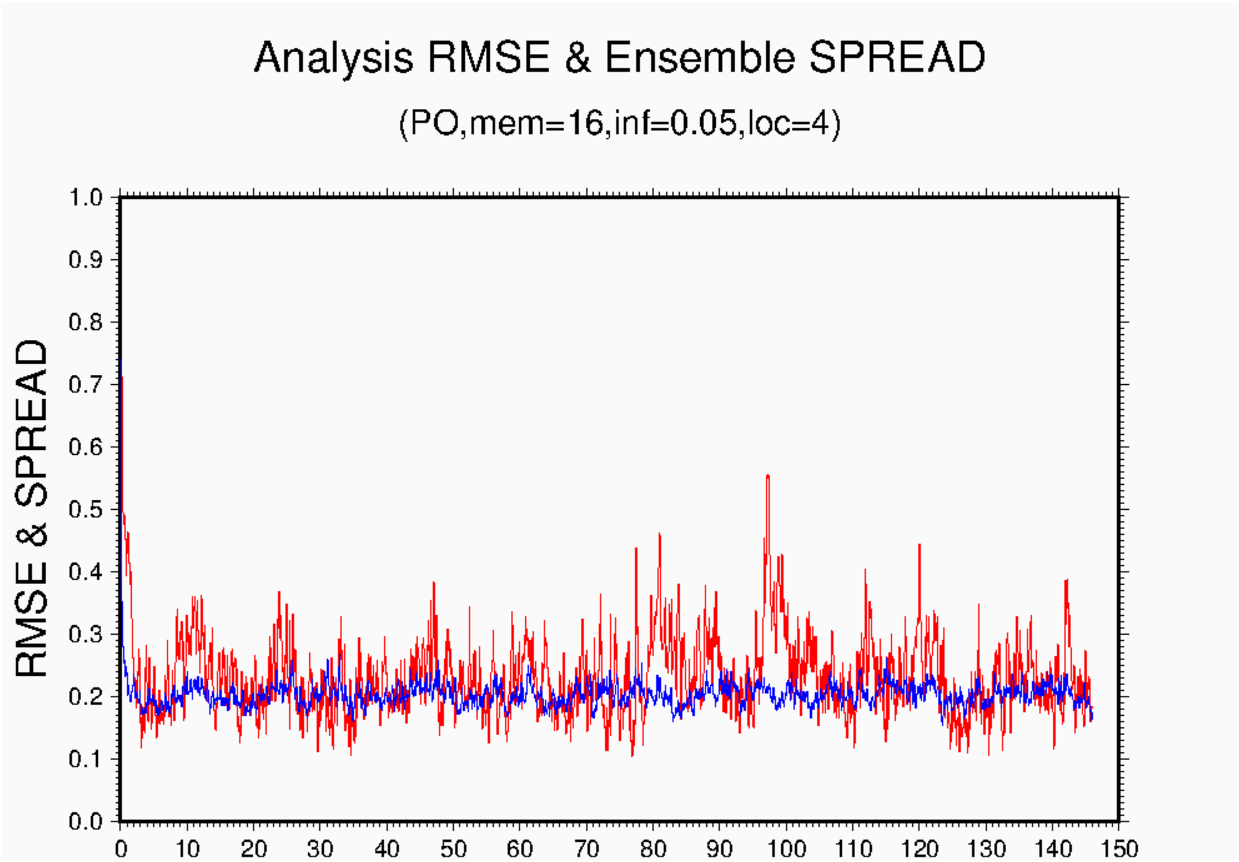


図 5: データ同化手法は PO-EnKF, アンサンブル数 16, 膨張係数 0.05, 局所化スケール 4 の時の解析 RMSE(赤線) と解析 SPREAD(青線) の時系列図. 横軸は時間, 縦軸は解析 RMSE と解析 SPREAD の値。

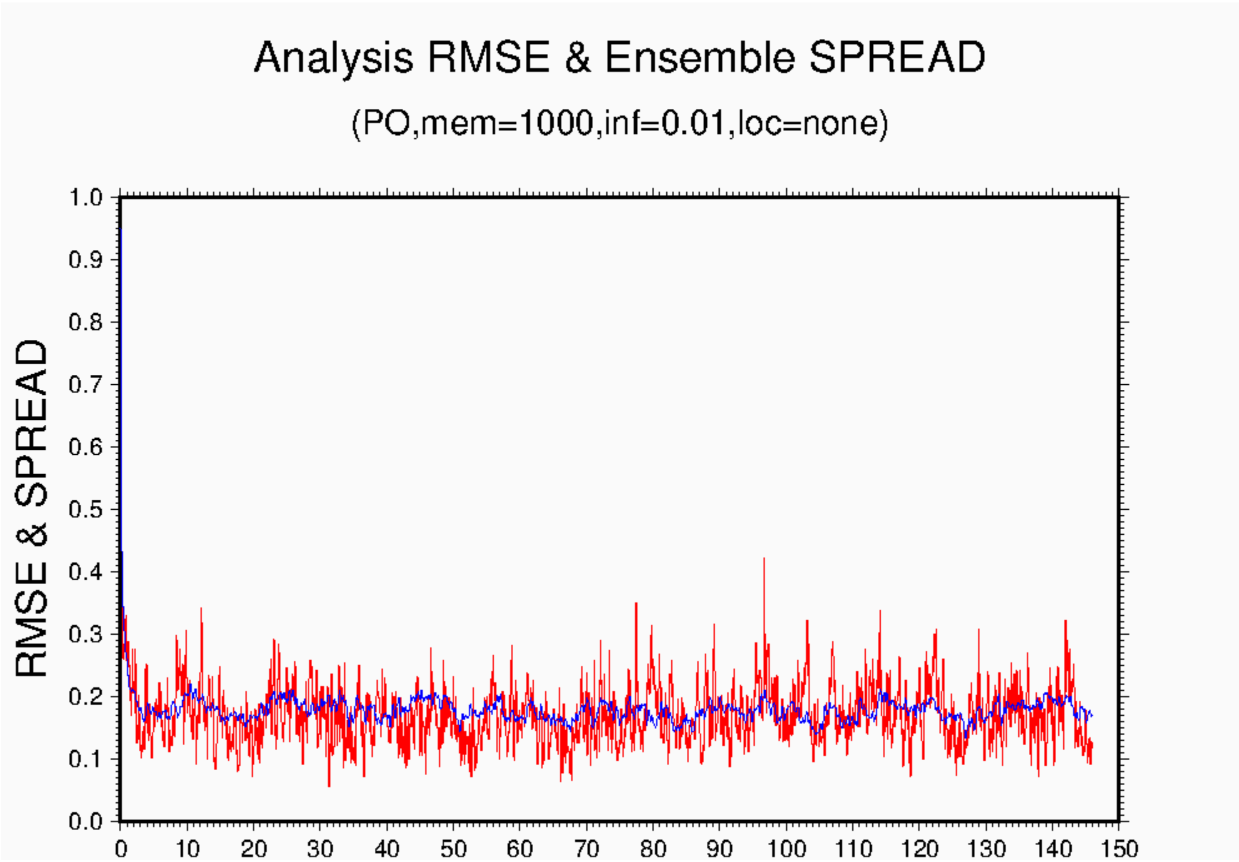


図 6: データ同化手法は PO-EnKF, アンサンブル数 1000, 膨張係数 0.01, 局所化を行わない場合の解析 RMSE(赤線) と解析 SPREAD(青線) の時系列図. 横軸は時間, 縦軸は解析 RMSE と解析 SPREAD の値。

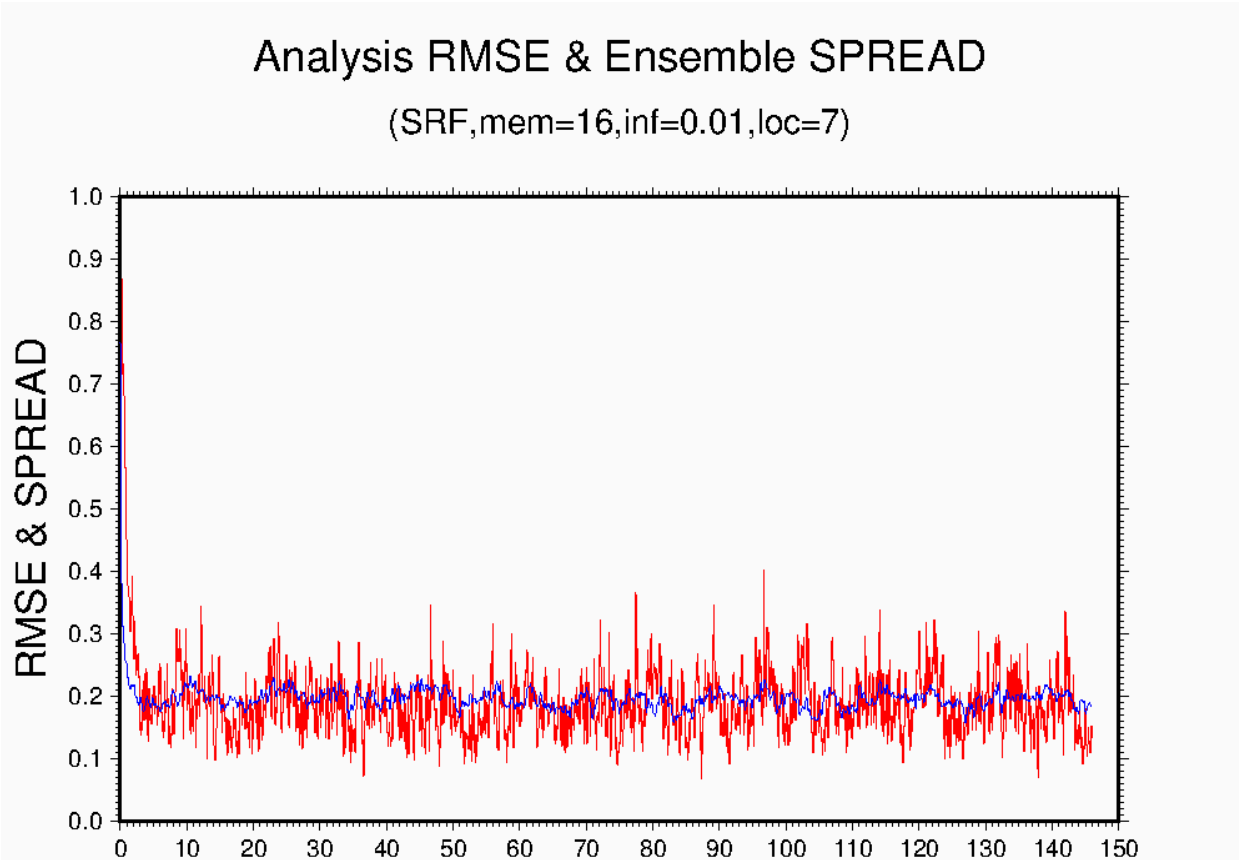


図 7: データ同化手法は Serial EnSRF, アンサンブル数 16, 膨張係数 0.01, 局所化スケール 7 の時の解析 RMSE(赤線) と解析 SPREAD(青線) の時系列図. 横軸は時間, 縦軸は解析 RMSE と解析 SPREAD の値。

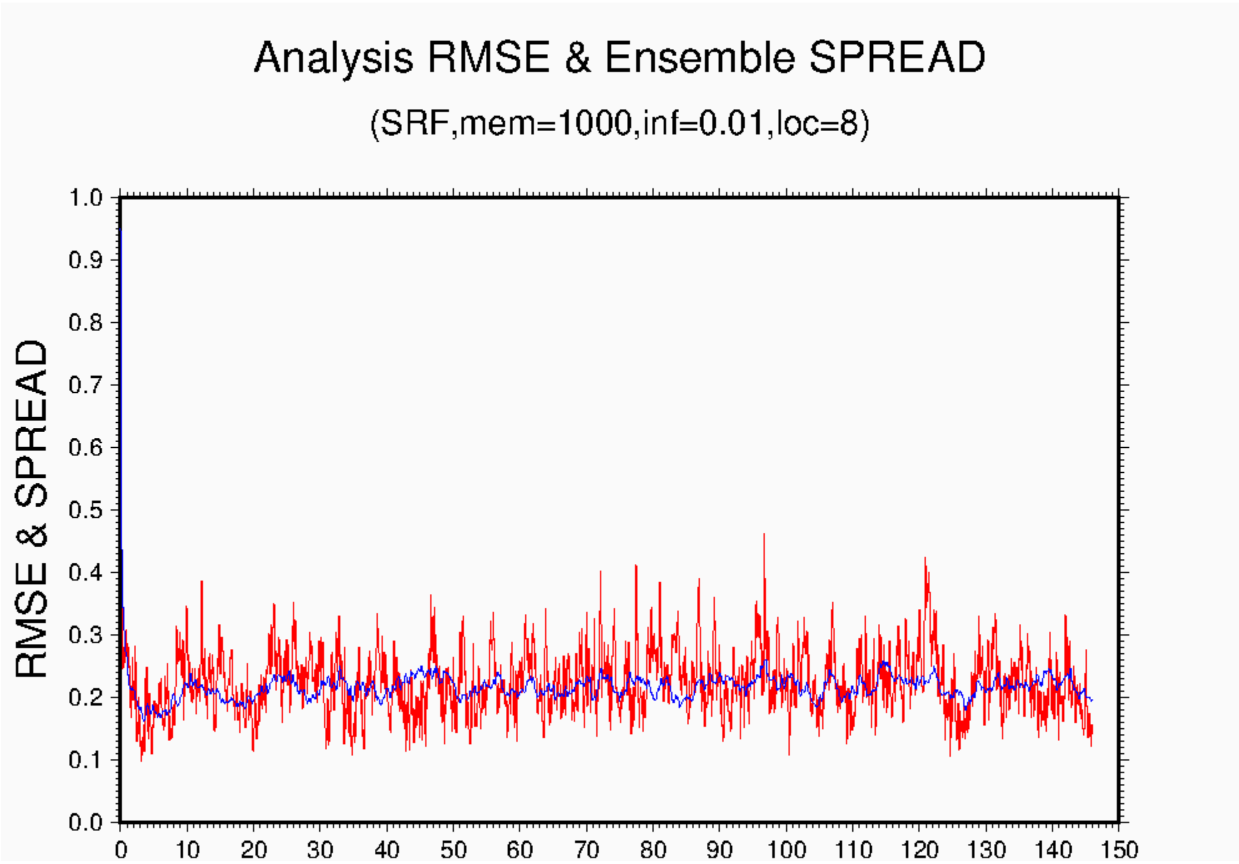


図 8: データ同化手法は Serial EnSRF, アンサンブル数 1000, 膨張係数 0.01, 局所化スケール 8 の時の解析 RMSE(赤線) と解析 SPREAD(青線) の時系列図. 横軸は時間, 縦軸は解析 RMSE と解析 SPREAD の値。

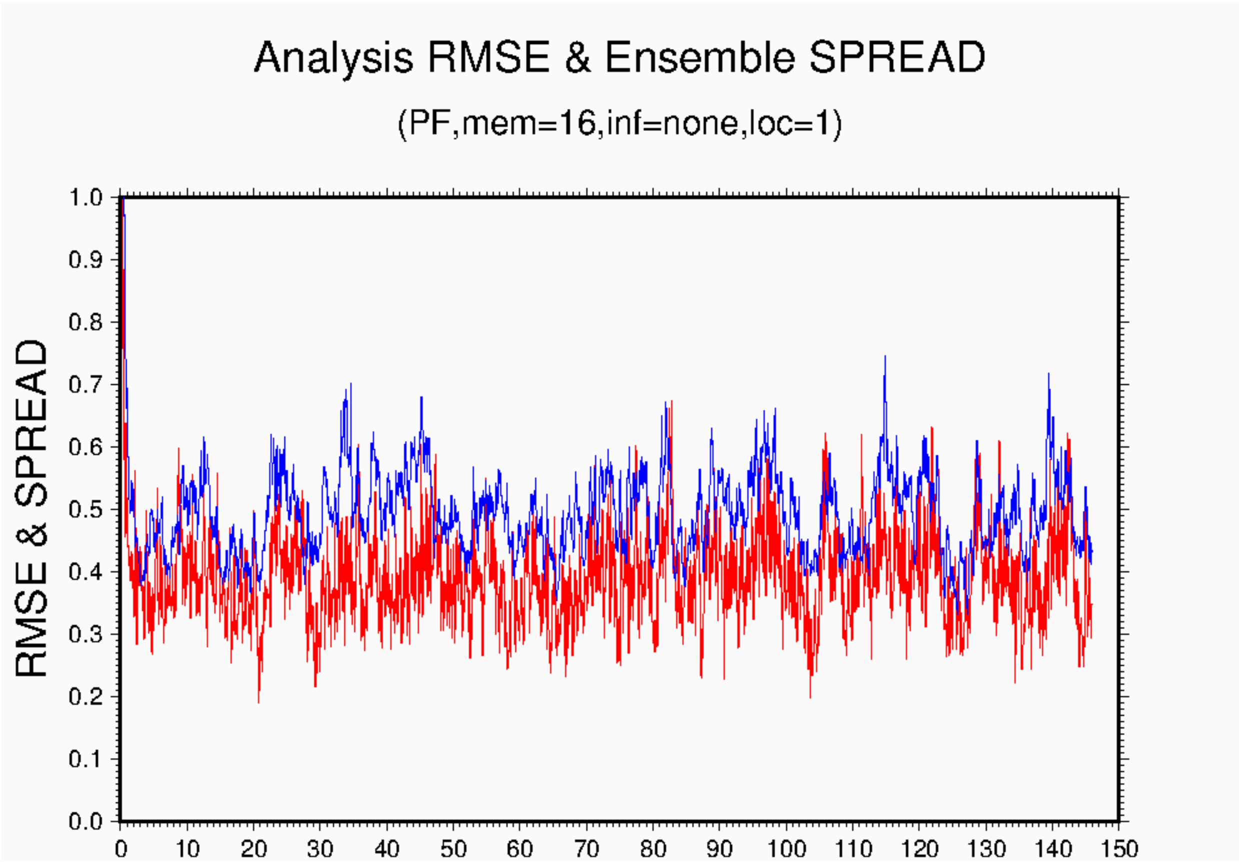


図 9: データ同化手法は PF, アンサンブル数 16, 局所化スケール 1 の時の解析 RMSE(赤線) と解析 SPREAD(青線) の時系列図. 横軸は時間, 縦軸は解析 RMSE と解析 SPREAD の値。

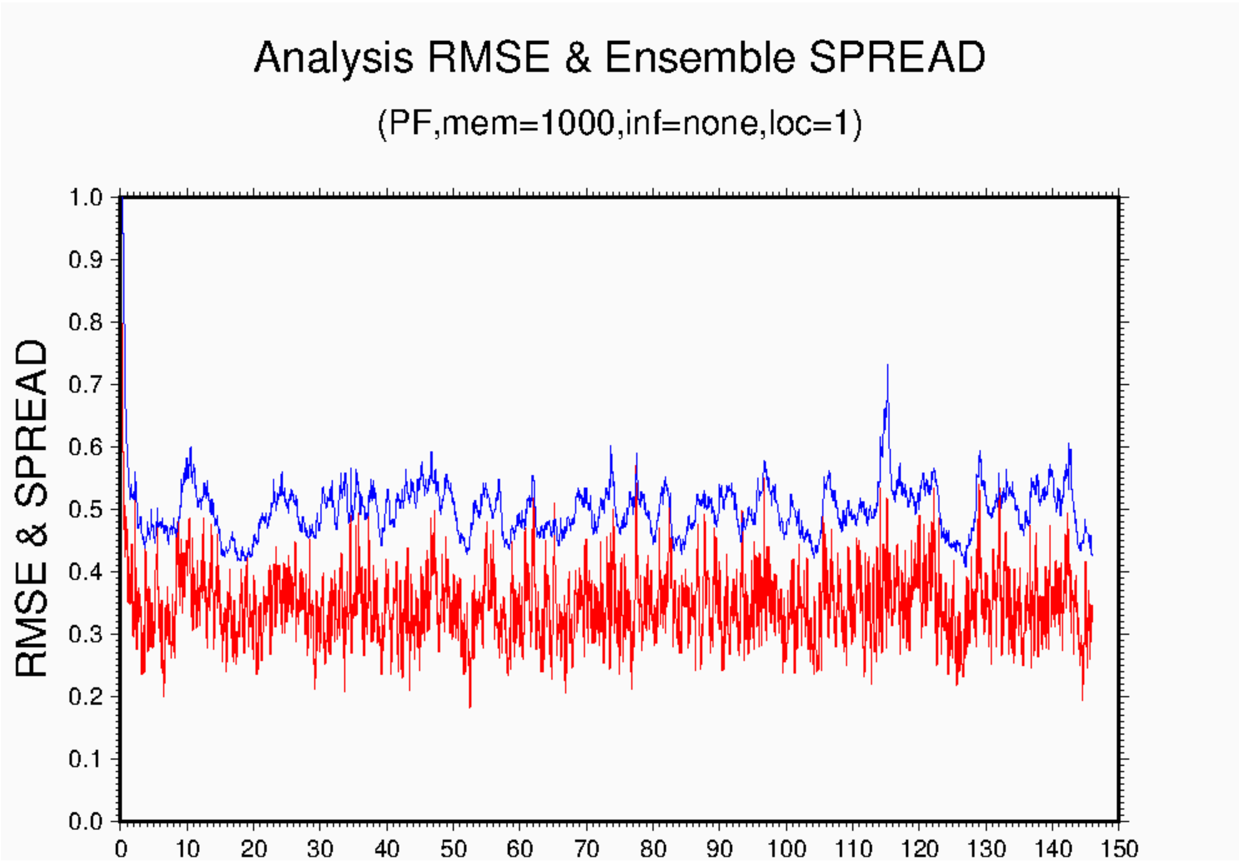


図 10: データ同化手法は PF, アンサンブル数 1000, 局所化スケール 1 の時の解析 RMSE(赤線) と解析 SPREAD(青線) の時系列図. 横軸は時間, 縦軸は解析 RMSE と解析 SPREAD の値。

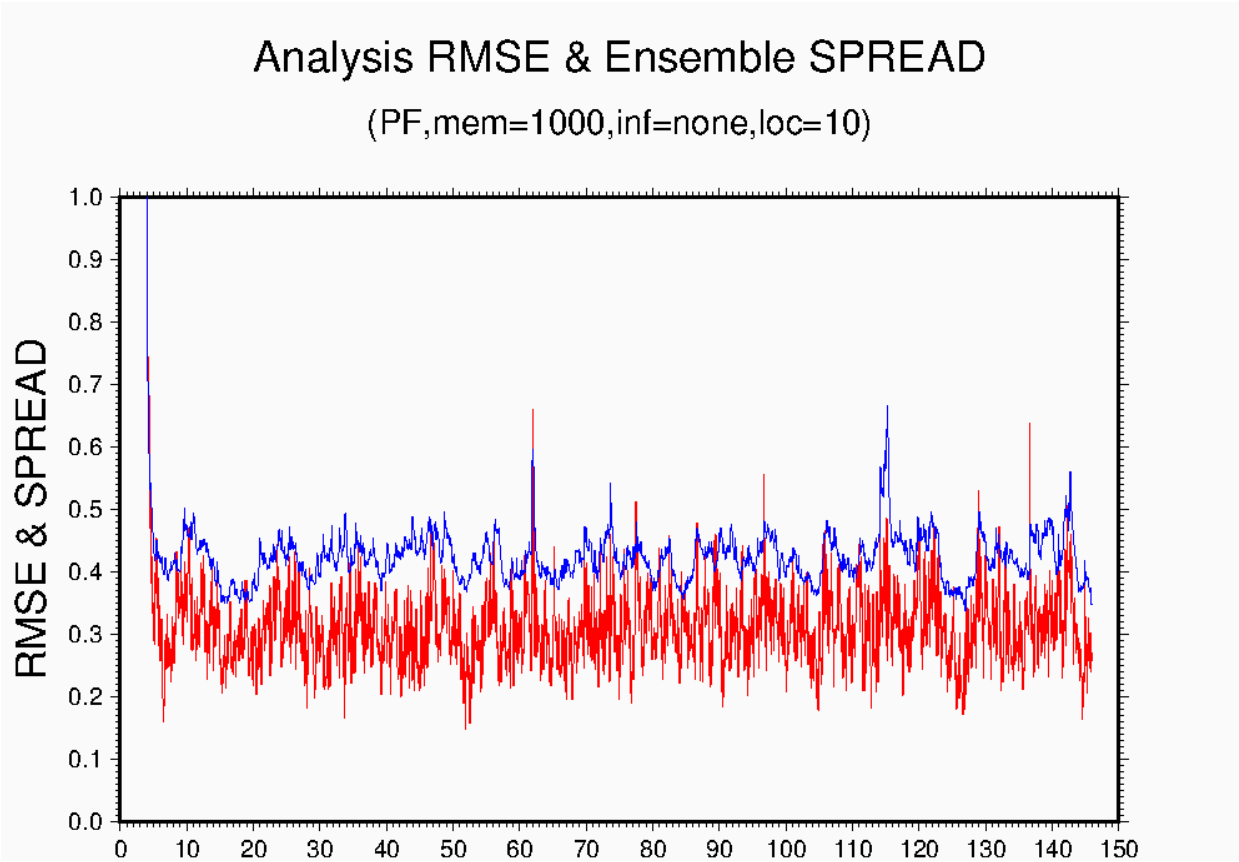


図 11: データ同化手法は PF, アンサンブル数 1000, 局所化スケール 10 の時の解析 RMSE(赤線) と解析 SPREAD(青線) の時系列図. 横軸は時間, 縦軸は解析 RMSE と解析 SPREAD の値。

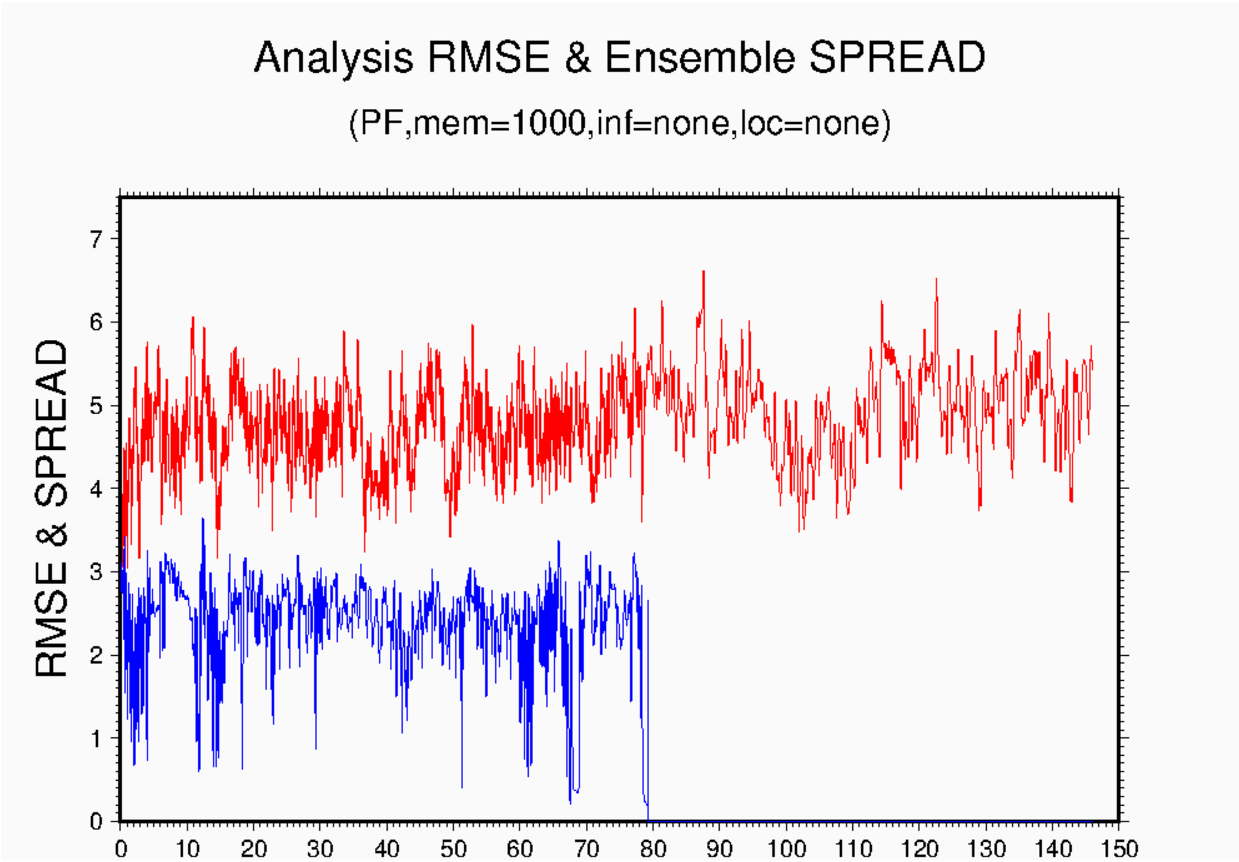


図 12: データ同化手法は PF, アンサンブル数 1000, 局所化を行わない場合の解析 RMSE(赤線) と解析 SPREAD(青線) の時系列図. 横軸は時間, 縦軸は解析 RMSE と解析 SPREAD の値。

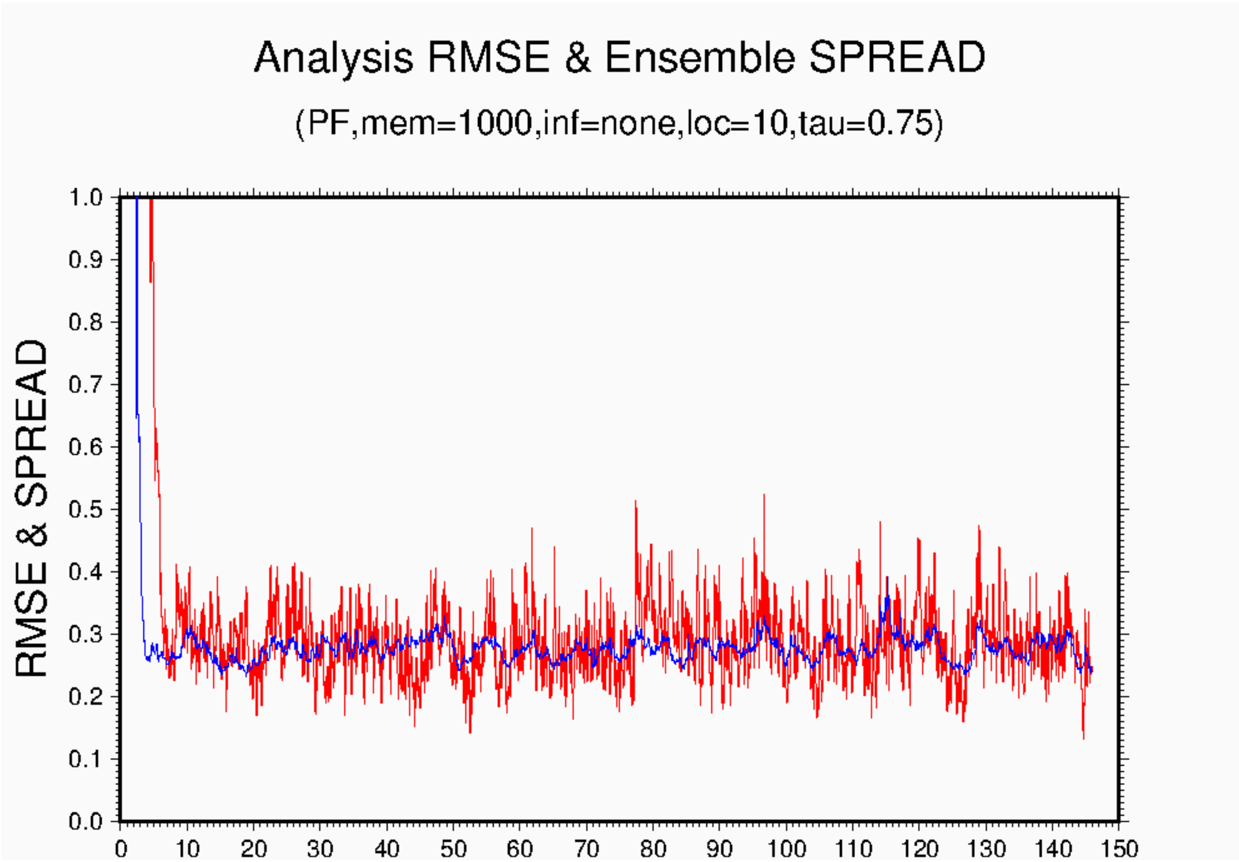


図 13: データ同化手法は PF, アンサンブル数 1000, 局所化スケール 10, 膨張係数 0.75 の時の解析 RMSE(赤線) と解析 SPREAD(青線) の時系列図. 横軸は時間, 縦軸は解析 RMSE と解析 SPREAD の値。

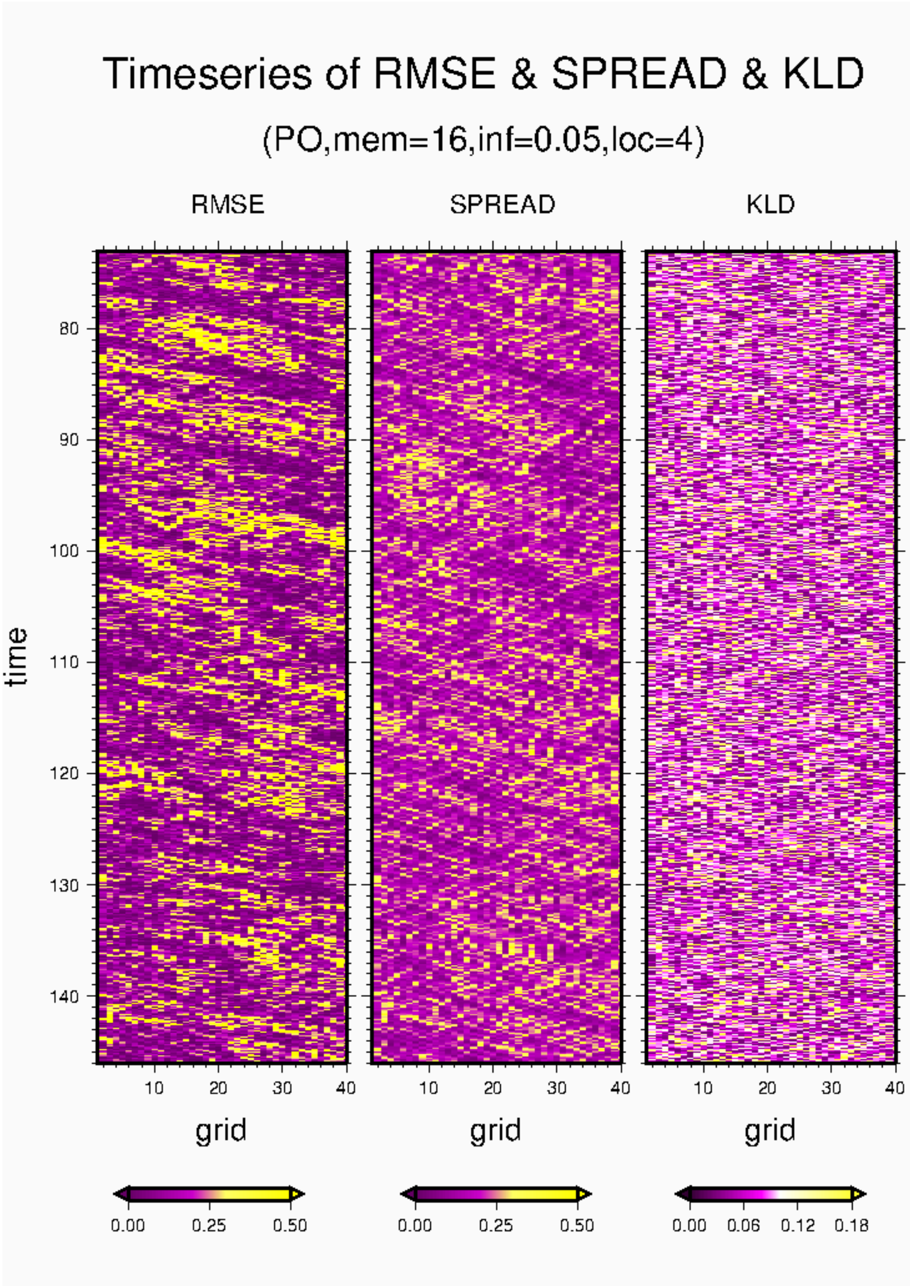


図 14: データ同化手法は PO-EnKF, アンサンブル数 16, 膨張係数 0.05, 局所化スケール 4 の時の解析 RMSE・解析 SPREAD・予報 KLD のホフメラー図.

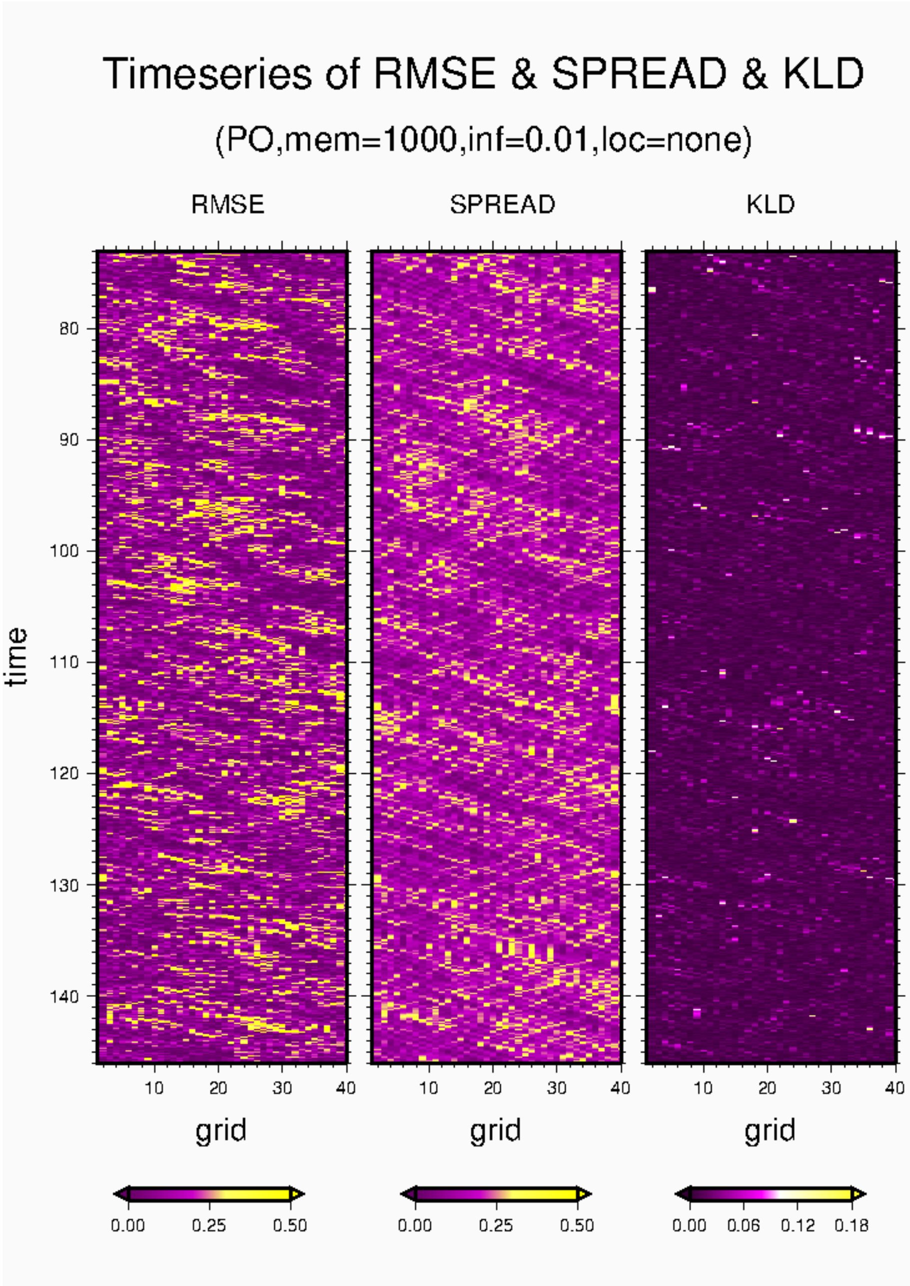


図 15: データ同化手法は PO-EnKF, アンサンブル数 1000, 膨張係数 0.01, 局所化を行わない場合の解析 RMSE・解析 SPREAD・予報 KLD のホフメラー図.

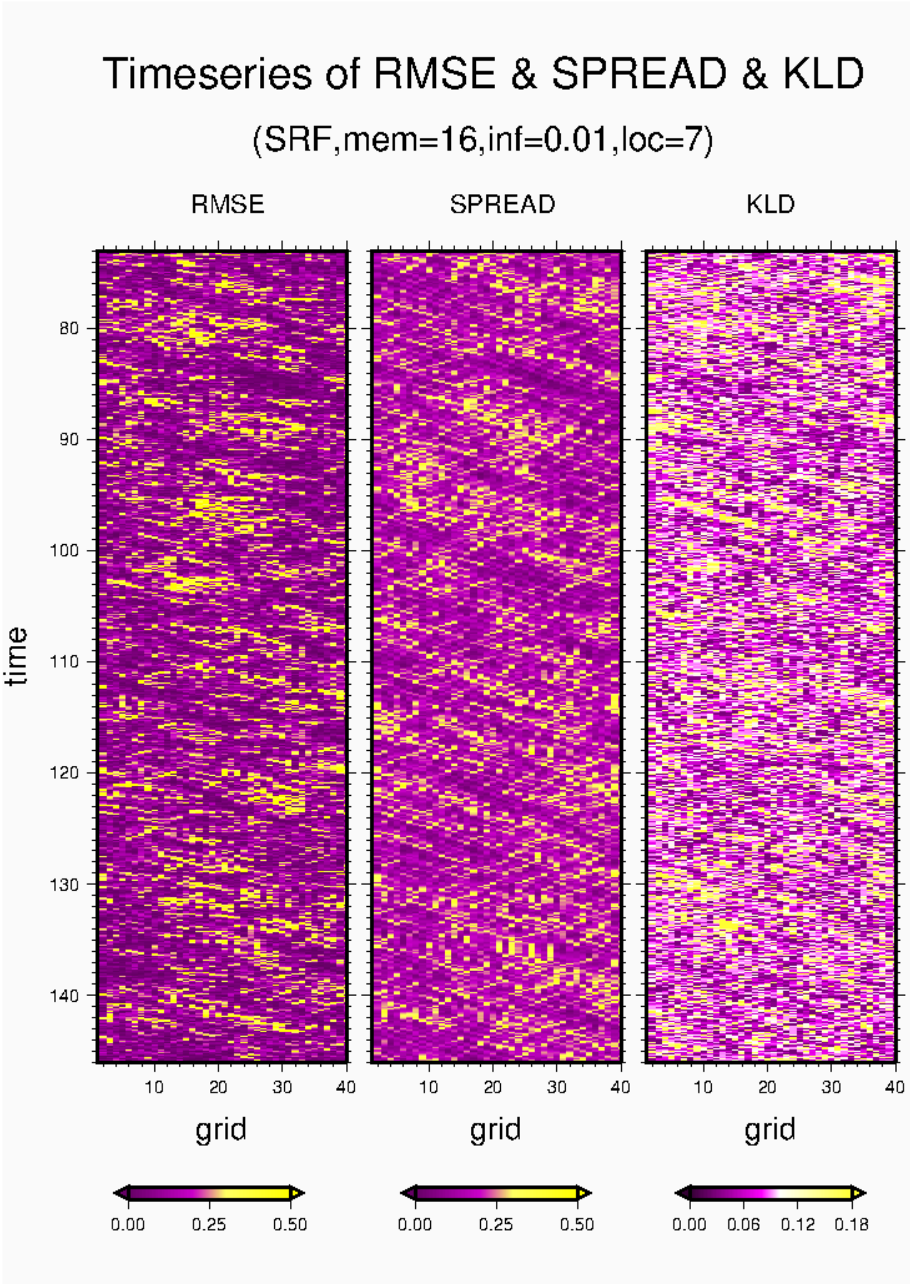


図 16: データ同化手法に Serial EnSRF, アンサンブル数 16, 膨張係数 0.01, 局所化スケール 7 の時の解析 RMSE・解析 SPREAD・予報 KLD のホフメラー図. 横軸は格子点番号で, 縦軸は時間。

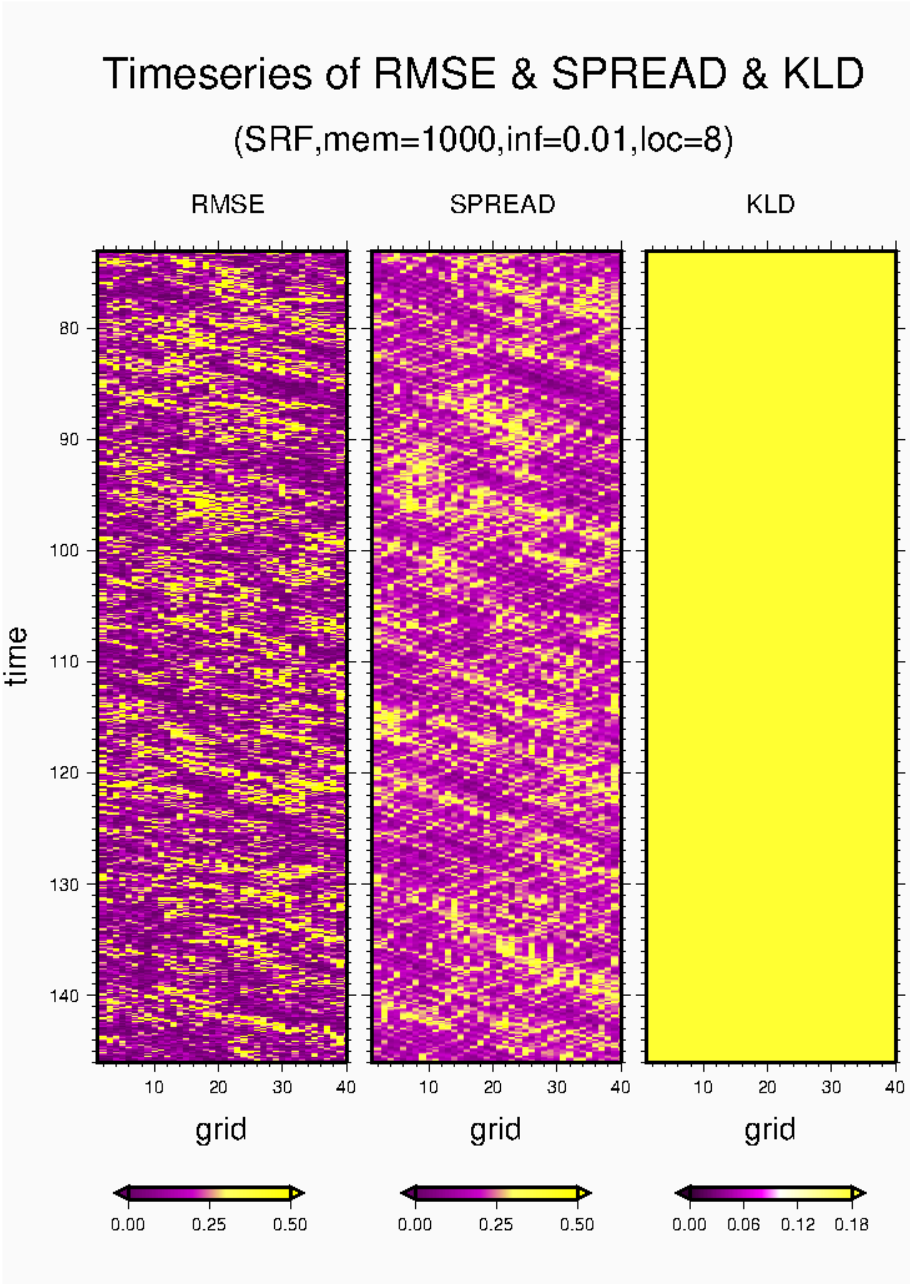


図 17: データ同化手法に Serial EnSRF, アンサンブル数 1000, 膨張係数 0.01, 局所化スケール 8 の時の解析 RMSE・解析 SPREAD・予報 KLD のホフメラー図. 横軸は格子点番号で, 縦軸は時間。

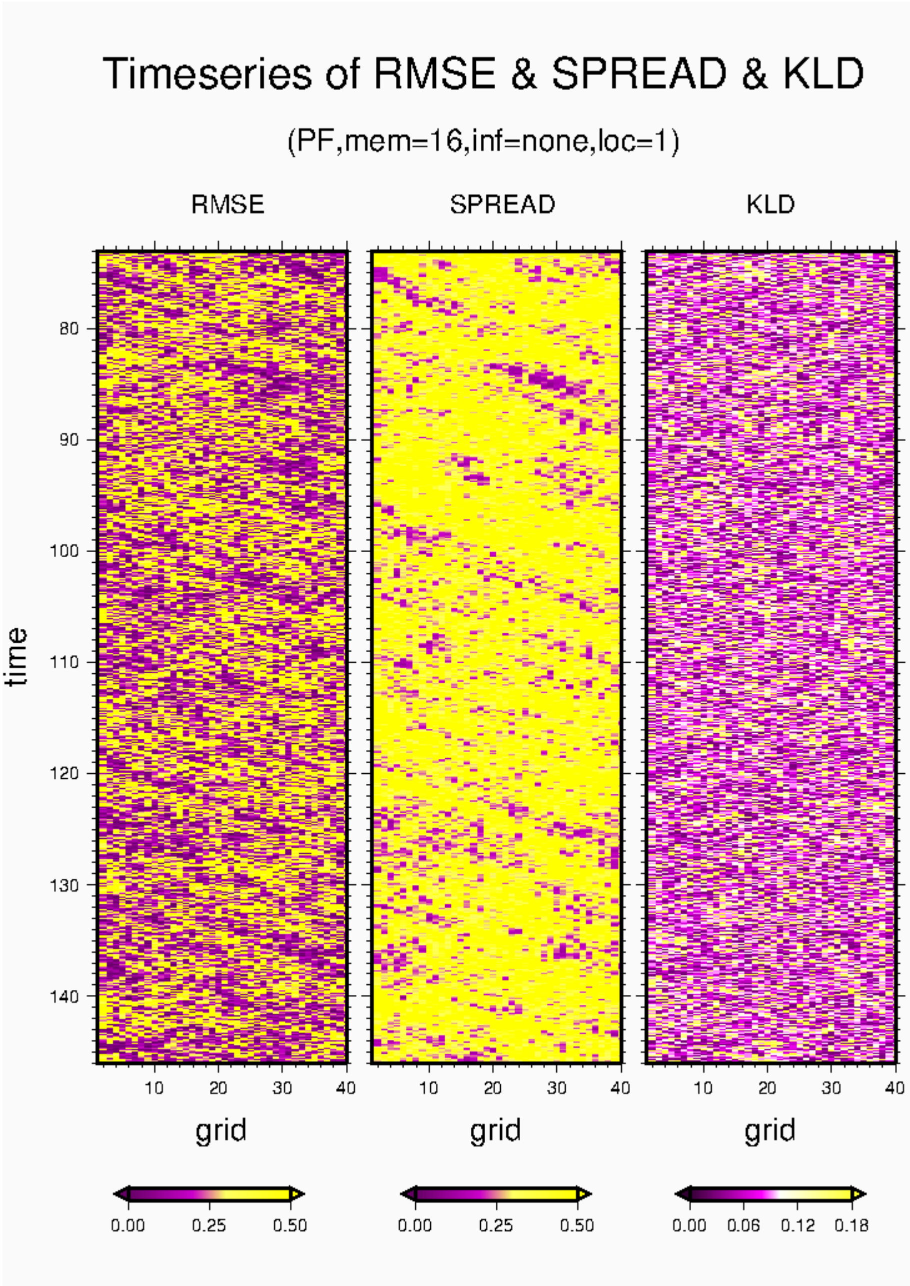


図 18: データ同化手法は PF, アンサンブル数 16, 局所化スケール 1 の時の解析 RMSE・解析 SPREAD・予報 KLD のホフメラー図.

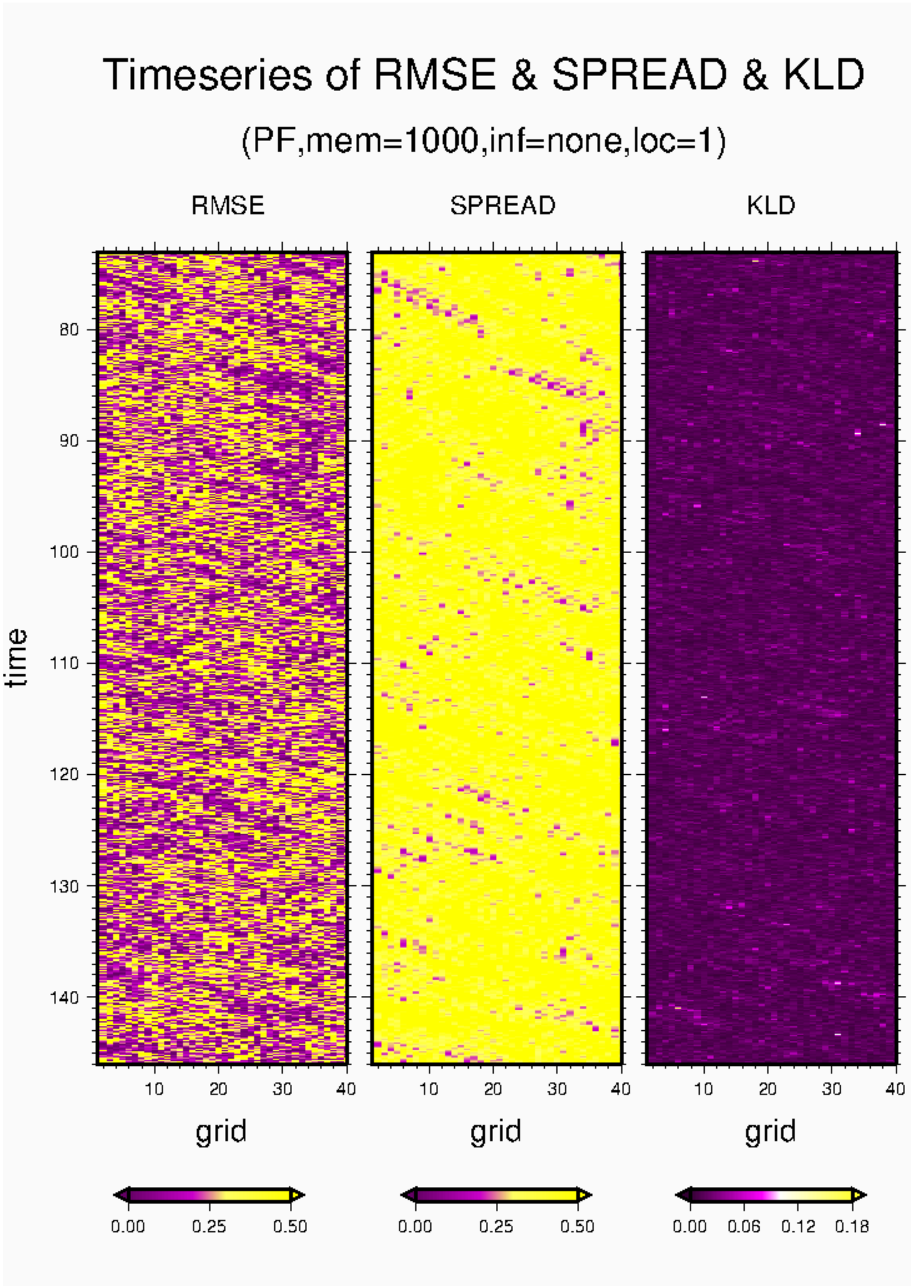


図 19: データ同化手法は PF, アンサンブル数 1000, 局所化スケール 1 の時の解析 RMSE・解析 SPREAD・予報 KLD のホフメラー図.

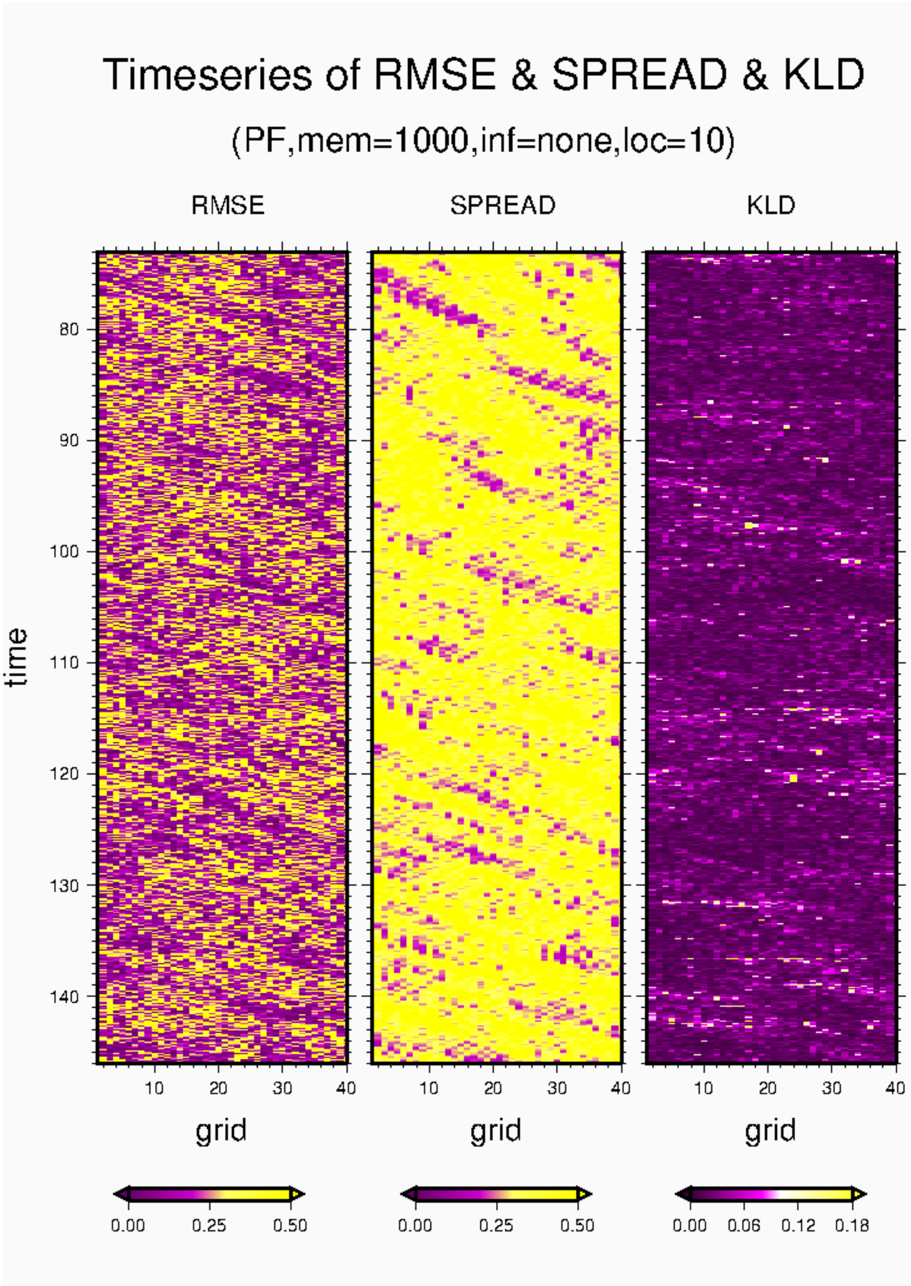


図 20: データ同化手法は PF, アンサンブル数 1000, 局所化スケール 10 の時の解析 RMSE・解析 SPREAD・予報 KLD のホフメラー図.

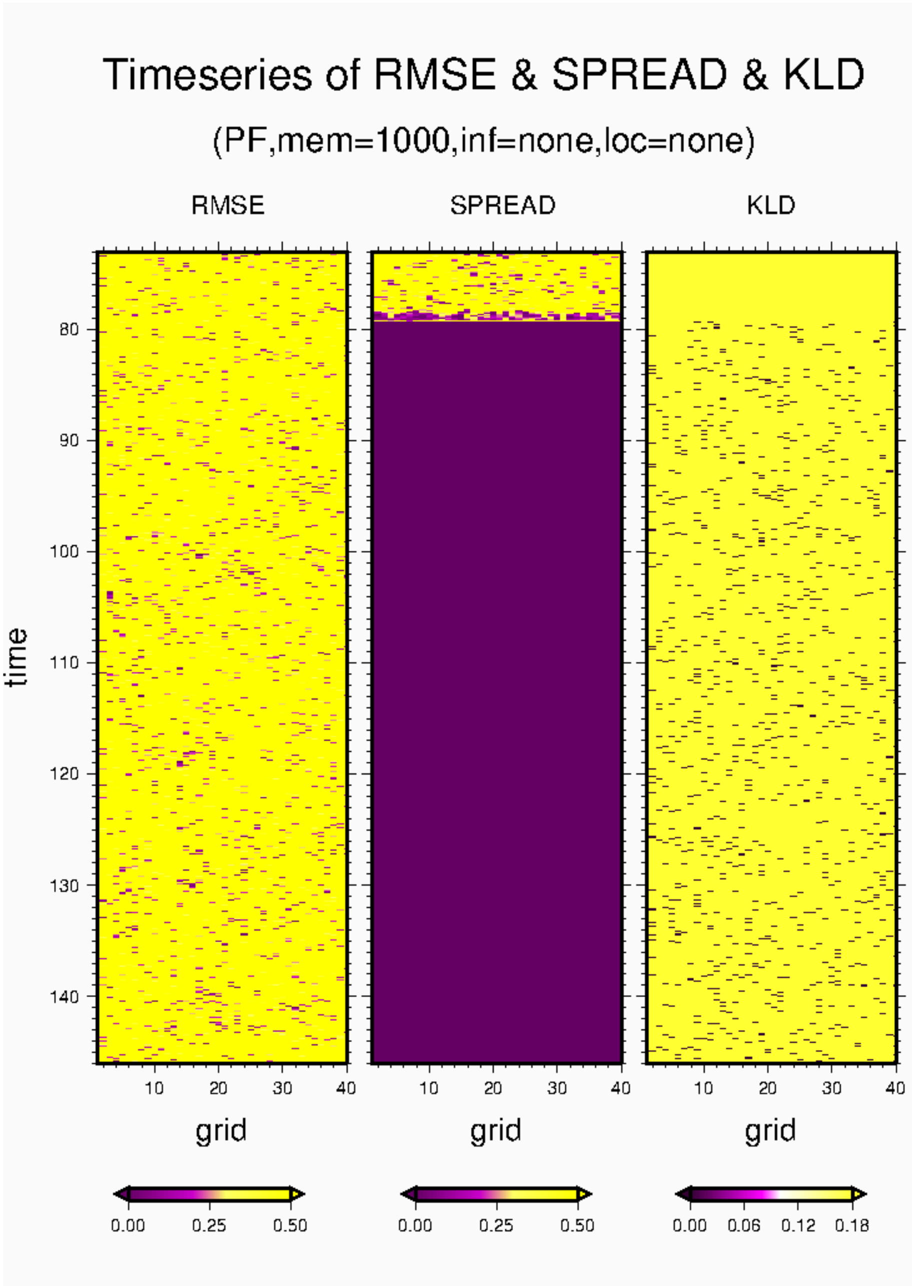


図 21: データ同化手法は PF, アンサンブル数 1000, 局所化を行わない場合の解析 RMSE・解析 SPREAD・予報 KLD のホフメラー図.

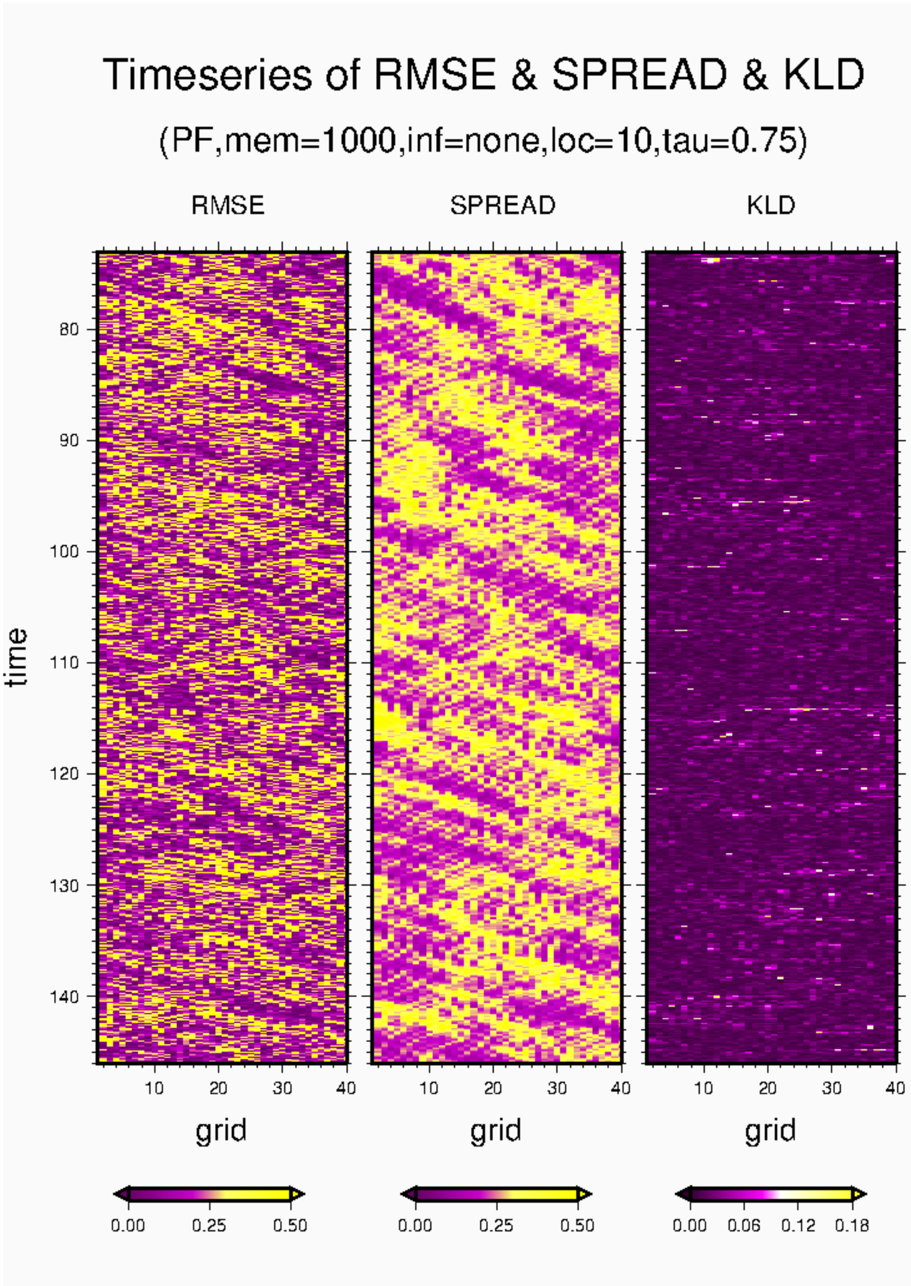


図 22: データ同化手法は PF, アンサンブル数 1000, 局所化スケール 10、膨張係数 0.75 の時の解析 RMSE・解析 SPREAD・予報 KLD のホフメラー図.

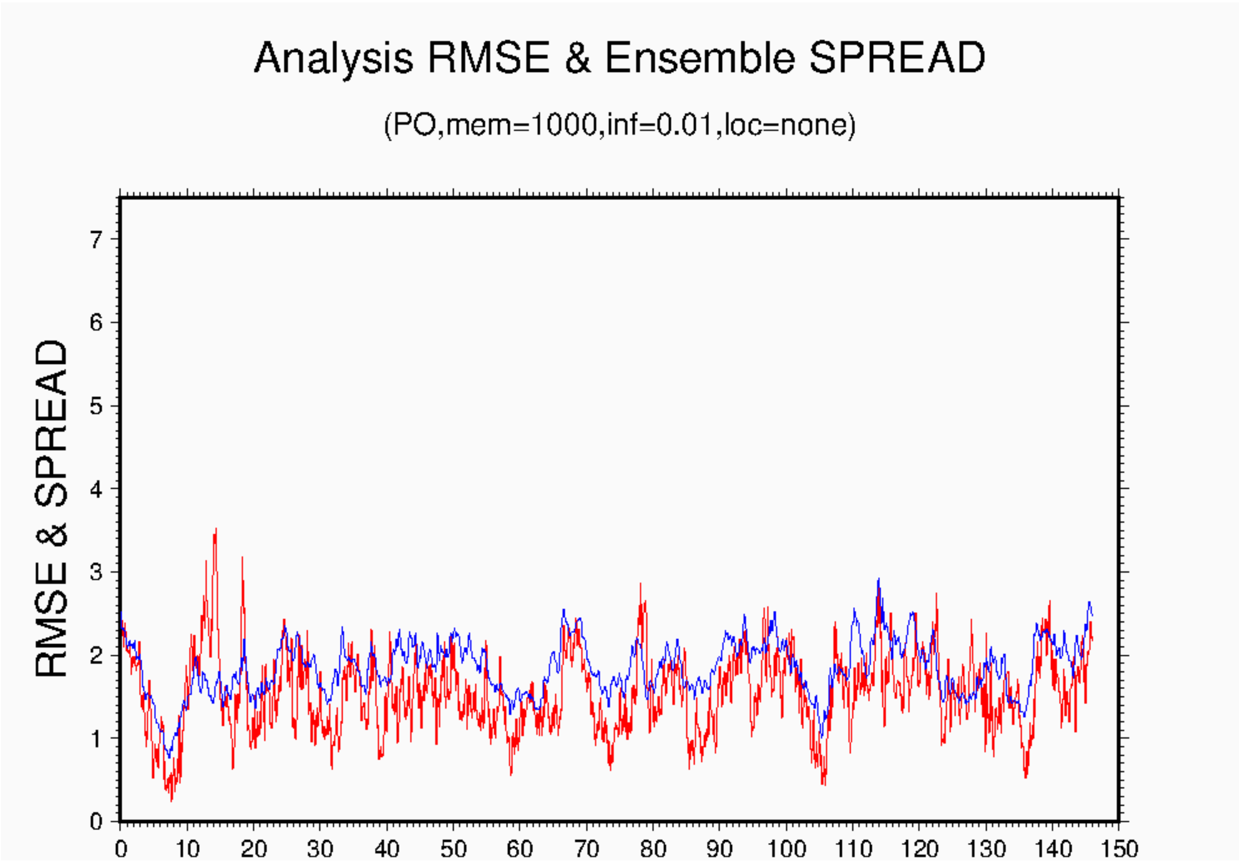


図 23: 観測点に疎密を与えた場合の解析精度の時系列図。がデータ同化手法は PO-EnKF, アンサンブル数 1000, 膨張係数 0.01, 局所化を行わない場合の解析 RMSE(赤線) と解析 SPREAD(青線) の時系列図. 横軸は時間, 縦軸は解析 RMSE と解析 SPREAD の値。

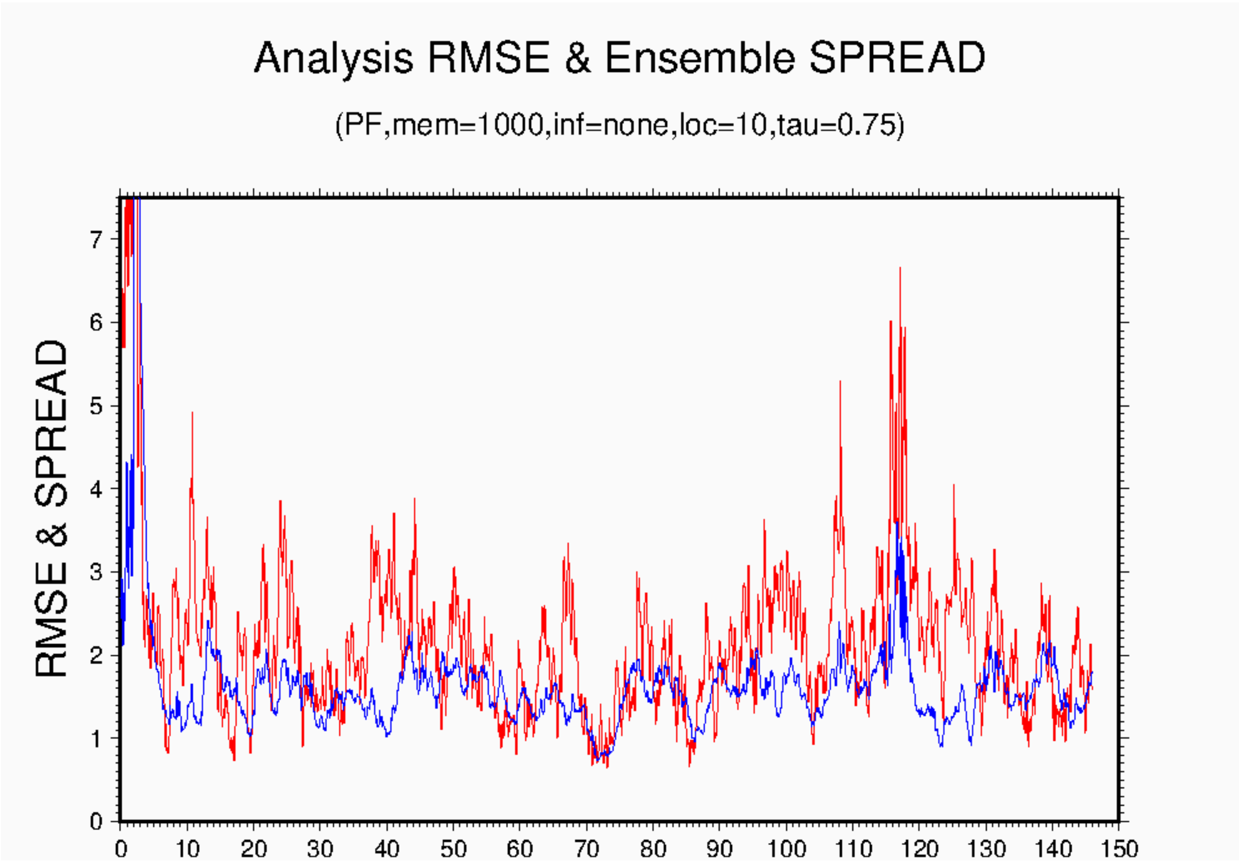


図 24: 観測点に疎密を与えた場合の解析精度の時系列図。データ同化手法は PF, アンサンブル数 1000, 局所化スケール 10, 膨張係数 0.75 の時の解析 RMSE(赤線) と解析 SPREAD(青線) の時系列図. 横軸は時間, 縦軸は解析 RMSE と解析 SPREAD の値。

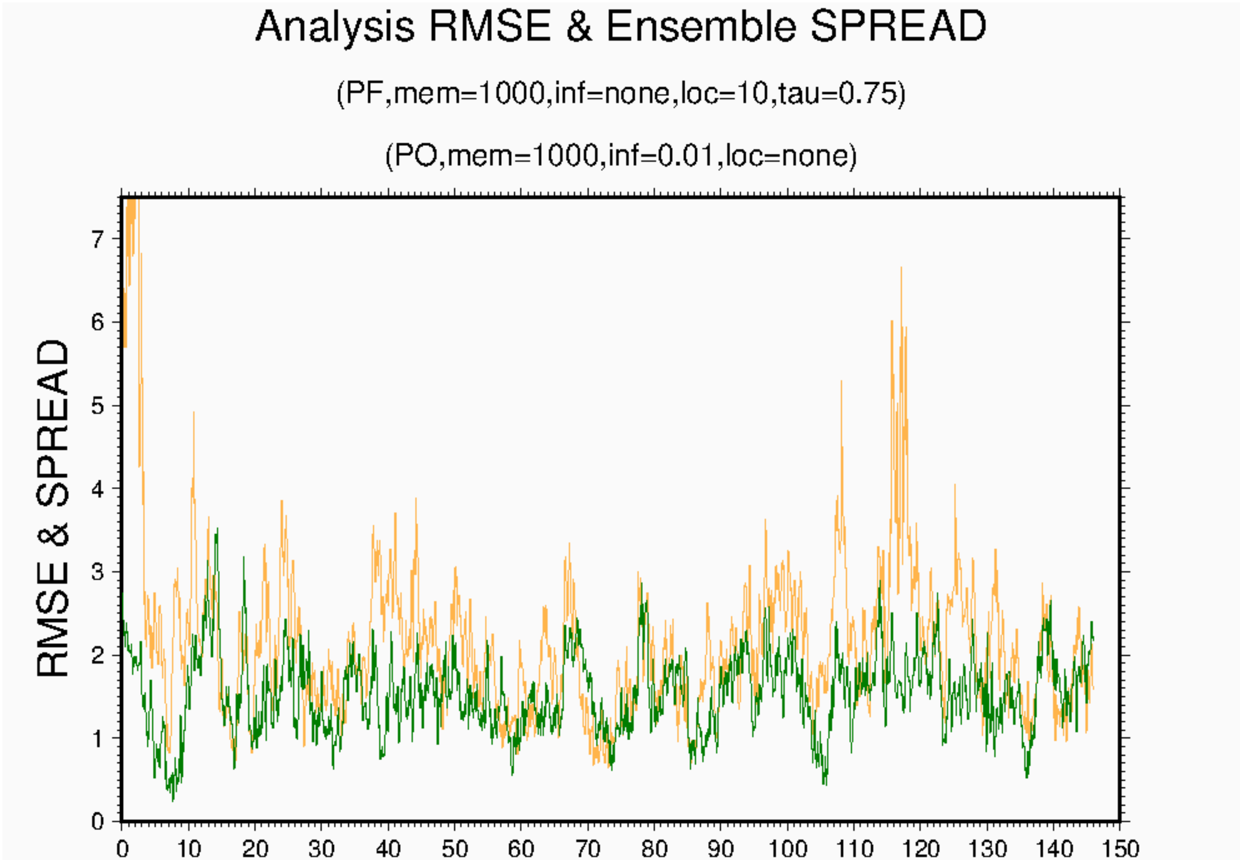


図 25: 観測点に疎密を与えた場合の解析精度の時系列図。PO-EnKF と PF の解析 RMSE のみ描画。PO-EnKF の解析 RMSE が橙線、PF の解析 RMSE が緑線、横軸は時間, 縦軸は解析 RMSE の値。

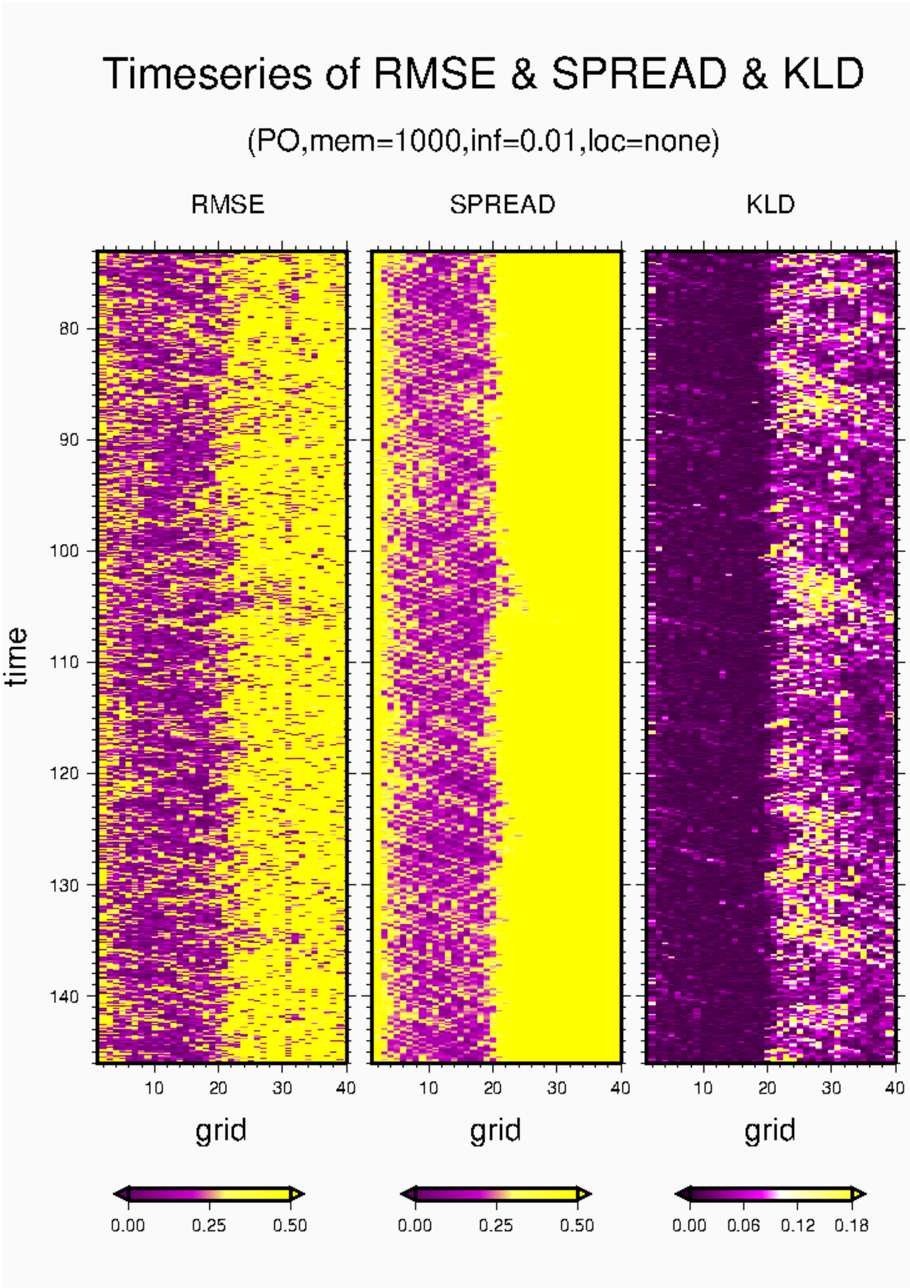


図 26: 観測点に疎密を与えた場合の解析精度のホフメラー図。データ同化手法は PO-EnKF, アンサンブル数 1000, 膨張係数 0.01, 局所化を行わない場合の解析 RMSE・解析 SPREAD・予報 KLD のホフメラー図.

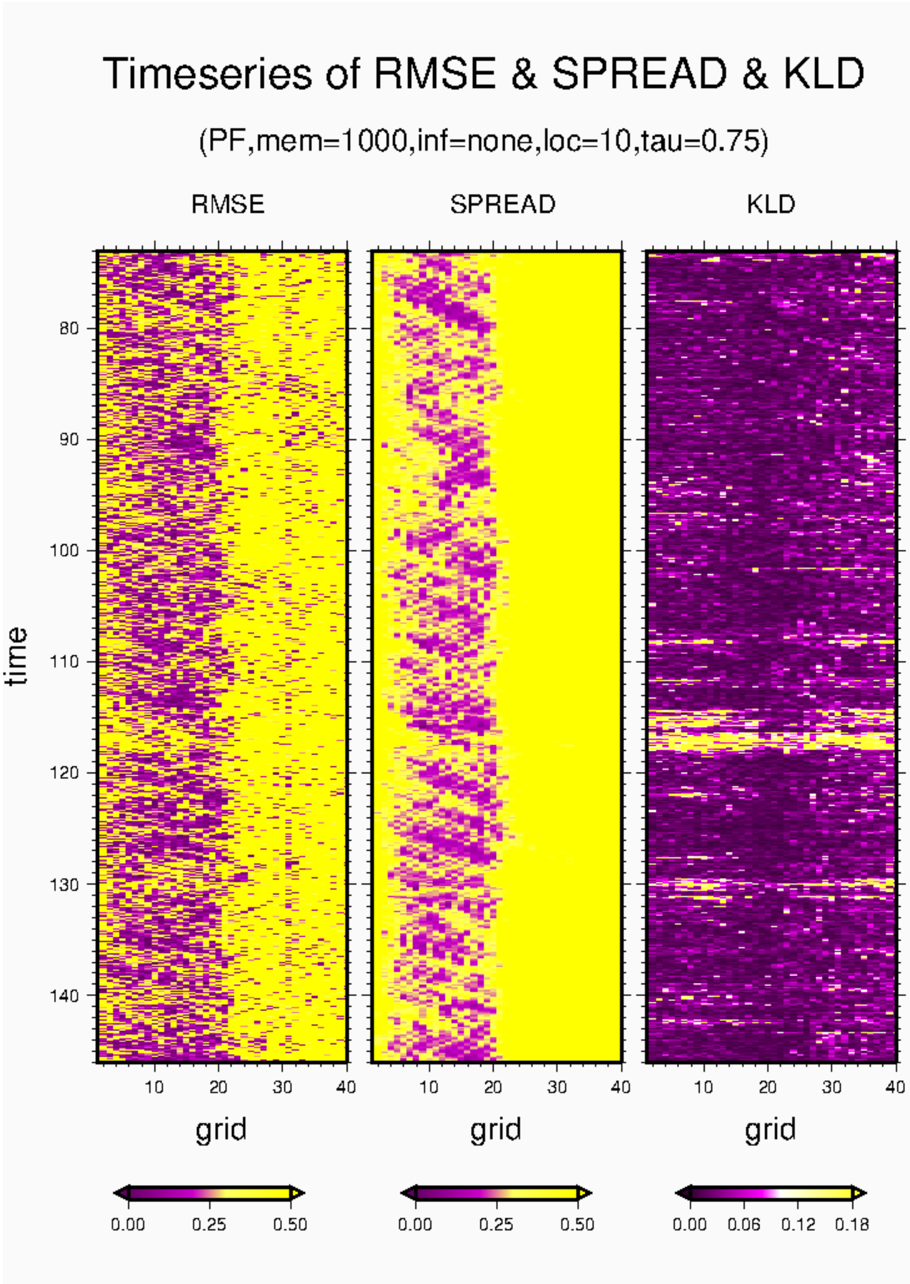


図 27: 観測点に疎密を与えた場合の解析精度のホフメラー図。データ同化手法は PF, アンサンブル数 1000, 局所化スケール 10、膨張係数 0.75 の時の解析 RMSE・解析 SPREAD・予報 KLD のホフメラー図.