今後の黒潮と周辺海域の
国際共同観測を考える

海洋変動予測モデルのアンサンブル化による観測との協働可能性について

宮　澤　泰　正
青　木　邦　弘

Feasibility of co-working between observation and ocean forecasting models by ensemble

Yasumasa Miyazawa・Kunihiro Aoki

みやざわ やすまさ：海洋研究開発機構
　アプリケーションラボ
あおき くにひろ　：海洋研究開発機構
　アプリケーションラボ

私たちは現在，海洋変動予測モデルのアンサンブル化の取り組みを進めている．アンサンブルの応用のひとつとして，モデルと観測がより密接に連携できるように開発を進めているアンサンブル感度解析について述べる．

1. はじめに

これまで，私たちは海洋循環モデルへのデータ同化による黒潮流路変動の予測実験を行ってきた．最近は，アンサンブルを活用して，黒潮流路変動の確率予測手法（Aoki *et al*. 2020a，改訂中）や，観測データの時間情報を活用する四次元変分法（Miyazawa *et al*. 2020，Ocean Dynamics, 受理済）の開発を進めている．今後，計算資源をさらに充実させることによってアンサンブルを用いた様々な応用解析が可能になると期待している．計算資源は一方でダウンスケーリング，高解像度化による活用がありうるが，解像度を変えずにアンサンブルを生成するという活用の方向性もある．アンサンブルを用いた有力な応用解析として，観測とモデル予測をより密接に連携させる（協働させる）ことが可能となるアンサンブル感度解析がある．以下ではアンサンブル感度解析の基本的な枠組と，観測との関連について述べる．

2. アンサンブルによる初期値感度解析

数値予報モデルは，ある時刻の状態から次の時刻の状態を予報する．これは，一般に

$$x_{t+\Delta t} = N(x_t) \quad (1)$$

という形の発展方程式で表せる．ここで，$x_t \in R^n$ は時刻 t における状態，Δt は時間間隔，N は非線形を含む演算子である．今，各時刻の状態からの摂動 δx を考えると，第一近似で，

$$\delta x_{t+\Delta t} = M\delta x_t \quad (2)$$

が成り立つ．ここで，$M \equiv \partial N / \partial x_t$ は $n \times n$ 行列をなす線形演算子である．(2) は，ある時刻の摂動（以降，初期摂動と呼ぶ）の Δt 時間後の線形発展を表している．本稿では，δx_t および $\delta x_{t+\Delta t}$ を，

それぞれ，y および z と表し，(2) を

$$z = My \tag{3}$$

の形で表す．また，以下では，特に断らない限り，初期時刻を $t = 0$ とし，Δt 後の時刻を t と改める．

初期値感度解析の目的は，「ある時刻における状態のある特徴が，時間を遡った前の時刻におけるどのような状態に起因するかを探る」ことにある．その”特徴”は，ひとつのスカラー量で定義され，これを目的変数，あるいは，評価関数，と呼ぶ．たとえば，黒潮流量変動に注目するなら，注目領域での流量の空間平均値が目的変数となる．目的変数を $\Theta(x_t)$ で表すと，状態に対するその感度は，変分 $\theta \equiv \delta\Theta$ で定義される．したがって，目的変数の感度は具体的には，

$$\theta = \left[\frac{\partial\Theta}{\partial x_t}\right]^T z = \left[\frac{\partial\Theta}{\partial x_0}\right]^T y \tag{4}$$

と表される．二つ目の等式は x_t から x_0 への座標変換である．上式中央の $\partial\Theta/\partial x_t$ は z に対する双対ベクトルをなし，これは，

$$M^T \frac{\partial\Theta}{\partial x_t} = \frac{\partial\Theta}{\partial x_0} \tag{5}$$

に従う．M^T はアジョイント演算子と呼ばれ，時間逆積分を表す．(3) と (5) は，(4) の座標変換を満たしていることが確認できる．

目的変数の最大感度を与える初期擾乱を得るには，そのノルムが 1($\|y\| \equiv y^t y = 1$) となる条件のもとで，次の λ を未定乗数とするラグランジュ関数

$$F(y, \lambda) = \theta + \lambda(1 - y^T y) \tag{6}$$

の y に対する停留条件を求めれば良い．すなわち，$\partial F/\partial y = 0$ より，

$$2\lambda y = M^T \frac{\partial\Theta}{\partial x_t} \tag{7}$$

を得る．ここで，(5) を用いた．

初期値感度解析のうち，目的変数を $\Theta = z^T z$ と選んだものは特に特異値解析と呼ばれることが多い．特異値解析の目的は，線形的に最も成長する初期擾乱を見つけ出すことにある．この場合，ラグランジュ関数は，(3) を用いて，

$$F(y, \lambda) = y^T M^T My + \lambda(1 - y^T y) \tag{8}$$

で表され，その y についての停留条件は，

$$M^T My = \lambda y \tag{9}$$

となる．これは，λ と y をそれぞれ固有値と固有ベクトルとする対称行列 $M^T M$ に対する固有値問題である．固有値は固有ベクトルが持つ成長率を表す．従って，これを解いて得られる最大固有値に対する y は最大成長率を持つ初期擾乱である．また，固有値問題に対する双対性（伊藤と見延，2010[1]）から，上式は次のように表すこともできる．

$$My = \sqrt{\lambda} z \quad and \quad M^T My = \lambda y \tag{10}$$

$$M^T z = \sqrt{\lambda} y \quad and \quad MM^T z = \lambda z \tag{11}$$

これは M についての特異値問題の一つの組みを与えている．多くの論文や解説書では，この形式が用いられる（Fujii *et al*. 2008[2] など）．(10) の左側の式は (3) と等価である．ただし，(10) では，対称な形で表すために y を $\sqrt{\lambda}$ で規格化していることに注意する．ここまで，目的変数の最大感度を与える初期擾乱が (7) や (9) の形で求まることを示してきた．ただし，それはアジョイント演算子が前もって分かっていることが前提である．一般には，アジョイント演算子を得るには，様々な変数と関連づけられた非線形演算子から，数値不安定を起こさないような線形演算子を構成するというこの大変な作業を伴うため，上記の方法は必ずしも実用的ではない．この問題を回避するため，Enomoto *et al*.（2015）[3] はアンサンブルメンバーを用いた初期値感度解析（以降，アンサンブル感度解析）を提案している．

Enomoto *et al*.（2015）[3] に従えば，m 個のアンサンブルメンバーが作る初期摂動は，一つのベクトルの組みとして表される．

$$Y = (y_1, y_2, ..., y_m) \tag{12}$$

y_i がn次元ベクトルであるから，Y は $n \times m$ の行列をなす．同様に，初期時刻から一定時間経過した後の各メンバーの摂動を，$n \times m$ 行列

$$Z = (z_1, z_2, ..., z_m) \tag{13}$$

で表す．今求めたいのは，目的変数の感度を最大化する初期摂動であるが，それが Y のいずれかのベクトルで表せる保証はない．そこで，初期摂動を各メンバーのそれの線形結合で表す．

$$y = p_1 y_1 + p_2 y_2 + ... + p_m y_m = Yp \tag{14}$$

ここで，$p^T = (p_1, p_2, ..., p_m)^T$ である．初期摂動が線形発展する期間においては，(3) より，z も同様に，

$$z = p_1 z_1 + p_2 z_2 + ... + p_m z_m = Zp \tag{15}$$

と表せる．これら線形結合における p を決めることがアンサンブル感度解析の目的である．

アンサンブルモデルは，ある時刻における状態をアンサンブルメンバーの数だけ生成する．従って，状態に基づく目的変数もメンバー毎に $\Theta_i(x_t)$ と定義でき，ゆえに，その感度もメンバー毎に

$$\theta_i \equiv \left[\frac{\partial \Theta_i}{\partial x_t} \right]^T z \quad (i = 1, 2, ..., m) \tag{16}$$

と定義できる．これらの感度をひとまとめに $\theta = (\theta_1, \theta_2, ..., \theta_m)^T$ と表す．しかし，我々が知りたいのは個々のメンバーが持つ初期摂動と感度の組みではなく，これらのメンバーが従うシステム (3) に内在する初期摂動と感度の組みである．そこで，そのような感度を，各メンバーが持つ感度の線形結合

$$\theta_{en} \equiv p_1 \theta_1 + p_2 \theta_2 + ... + p_m \theta_m = \theta^T p \tag{17}$$

で定義する．それにより，前述の一般的な感度解析と同じ議論が可能となる．すなわち，ここで考えるべきラグランジュ関数は，(6) に (14) と (17) を代入したものとなる．

$$F(p, \lambda) = \theta^T p + \lambda (1 - p^T Y^T Y p) \tag{18}$$

ただし，y が p に変換されていることに注意する．これの p についての停留条件は，$\partial F / \partial p = 0$ より，

$$2\lambda p = (Y^T Y)^{-1} \theta \tag{19}$$

である．従って，目的変数の最大感度を与える初期擾乱は，上式に左から Y を乗じた，$2\lambda y = Y(Y^T Y)^{-1}\theta$ として得られる．

特異値解析の場合は目的変数に $\Theta_i = z_i^T z_i$ を選ぶ．このとき，ラグランジュ関数は

$$F(p, \lambda) \equiv p^T Z^T Z p - \lambda (1 - p^T Y^T Y p) \tag{20}$$

となり，p についての停留条件として，

$$(Y^T Y)^{-1} Z^T Z p = \lambda p \tag{21}$$

を得る．(21) は一般化固有値問題であり，これを解くと p が得られ，それに応じて (14 式) より y が求まる．

Enomoto *et al*.（2015）[3] も言及しているように，アンサンブル感度解析は必ずしも万能な手法ではない．アンサンブル感度解析が一般的な初期値感度解析と整合するのは，厳密には，初期摂動の部分空間が一致するとき，すなわち，$m = n$ のときに限る．しかし，システムの自由度がより十分小さいとみなせる場合，アンサンブルメンバーの初期値をうまく選べば，アンサンブル感度解析は有効な手法となり得る．たとえば，黒潮流路変動は，その平衡状態が大蛇行，非大蛇行接岸流路，および，非大蛇行離岸流路の三つが支配的なため，アンサンブル感度解析は有効と考えられる．

アジョイント (7) あるいはアンサンブル (19) による最大予測感度領域を有効に検出することができれば，事前にその海域を観測することにより予測精度を向上させることができると考えられる．このような機動的観測の可能性は，最近の無人観測手段の発達や，測器の小型化・軽量化・通信の

簡易化によって従来よりも実現しやすくなっているのではないかと思われる．今後，様々な観測研究者，あるいは，漁業者など，現場での予測結果の活用者と連携し，機動観測の可能性を追求していきたい．

3. アンサンブルによる背景誤差共分散の表現

モデルの成長モードをうまく表現しているようなアンサンブルメンバーが得られた場合，これを使ってデータ同化にとって重要な背景誤差共分散を推定することも意義深い．実際，アンサンブルカルマンフィルターが行っていることそのものである．しかし，アンサンブルメンバー数が少ない場合はサンプリング誤差によって見かけの誤差共分散が現れる場合がある．またモデルに系統的な誤差がある場合（蛇行しやすい等），アンサンブルスプレッドが著しく小さくなり（縮退），アンサンブルによる誤差共分散の表現は正確さを失ってしまう．このためアンサンブルカルマンフィルターの研究分野においては，サンプリング誤差を減らすための「共分散局所化」と，アンサンブルの縮退を防ぐための「共分散膨張」についての研究開発が精力的に進められている．

ここでは，アンサンブルメンバーによる共分散の表現を従来の経験的な共分散の表現と融合させていくための手続きについて，Yaremchuk *et al*.（2011）[4] に従って記述する．Evsenen and van Leeuwen（1996）[5] の式 (7) を用いれば，アンサンブルメンバーから得られる初期摂動の標本共分散行列（n 行 n 列）は

$$\frac{YY^T}{m-1} \tag{22}$$

と表現される．(22) の行列はアンサンブルのサンプリング誤差の影響を受け，悪条件となるので，逆行列を求める際の精度が落ちてしまう．そこで，特異値 S と左右特異ベクトル U，V を用いて行列 Y を特異値分解

$$Y = USV^T \tag{23}$$

したときの m 個の特異値のうち，絶対値がある程度大きいものの個数を m* として，共分散行列 (22) を再構成し，

$$P\Lambda_{m*}P^T \tag{24}$$

をアンサンブルによる背景誤差共分散の近似とする．ここで，$P=(e_1,e_2,...,e_{m*})$ は，m* 番目までの特異ベクトル U から成る n 行 m* 列行列である．Λ_{m*} は，特異値の二乗値 $s_1^2,s_2^2,...,s_{m*}^2$ を成分とする対角行列である．$P^TP=I_{m*}$（m* 行 m* 列の単位行列）となるが，P^TP （n 行 n 列）は単位行列とならないことに注意する．経験的な表現として用いる背景誤差共分散行列を B_0 とすると，$P_{\oplus}=I_n-PP^T$ として

$$B=\alpha^{-1}P\Lambda_{m*}P^T+\beta^{-1}(P_{\oplus}B_0P_{\oplus}^T)^+ \tag{25}$$

を融合させた背景誤差共分散行列が定義できる．この B を背景誤差共分散行列の融合行列，あるいは，融合表現と呼ぶ．ここで，上付きの「+」は Moore-Penrose 一般化逆行列を表わす．α,β は正値の係数であり，それぞれアンサンブルによる背景誤差共分散行列（上式の右辺第一項）と，経験的背景誤差共分散行列（上式の右辺第二項）の重みを表わす．(25) の逆行列は，

$$B^{-1}=\alpha P\Lambda_{m*}^{-1}P^T+\beta P_{\oplus}B_0P_{\oplus}^T \tag{26}$$

となる．$P^TP_{\oplus}=P_{\oplus}^TP=0$ より，融合行列 (25) およびその逆行列 (26) における右辺第一項と右辺第二項は互いに独立（直交）である．

固有ベクトルの次数 m* は，下記のベイジアン情報量基準 (BIC) を最小化する値を与えるものとする．

$$C(m*)=m*+\frac{N}{\ln N}\ln\sigma_{m*}^2\to\min_{m*} \tag{27}$$

ここで，N は観測データサンプルの大きさで，具体的には，観測データ取得の時間的な回数を表す．σ_{m*}^2 は，m* 個の固有ベクトルによる観測

データサンプルN個あたりの摂動場の表現に関する近似誤差（BIC における尤度に対応）であり，

$$\sigma_{m*}^2 = \overline{\sum_{k=1}^{K_j}\left[\sum_{i=1}^{m*} f_i^j e_i(r_k^j) - y(r_k^j)\right]^2 \Bigg/ \sum_{k=1}^{K_j} y(r_k^j)^2} \quad (28)$$

で定義される．r_k^j は，j 回目に得られた k 個目の観測値の位置を表し，K_j は j 回目に得られた観測値の総数を表す．$y(r_k^j)$ は位置 r_k^j における初期値の摂動であり，$f_k^j e_j(r_k^j)$ は，固有ベクトルによる摂動の近似であり，行列による近似 (23) から得られる．上付きのバーは，N サンプル全体での平均操作を意味する．

重みづけ係数 α,β は，アンサンブルによる表現と経験的な表現の相対的な重みづけを与えるとともに，予測誤差の絶対値を決める値なので，同化解析値 x^a と予測値 x^f の差（インクリメント）$dx = x^a - x^f$ と，観測値 o とこれに対応する予測値 Hx^f の差 $do = o - Hx^f$ の情報にもとづいて決める（H は観測演算子を表す）．具体的には，行列の対角成分（Diag）同士を比較する

$$Diag(P^T \delta x \delta x^T P) \sim Diag(\frac{1}{\alpha}\Lambda_{m*}) \quad (29)$$

という近似によって α を決める．β は，$Tr\left\{E(do^T do)\right\} = Tr\left\{HBH^T + R\right\}$ という関係から（Tr は行列の対角成分の和を，E は期待値，R は観測誤差共分散行列を表わす），

$$\beta = \frac{Tr\left\{H\left[P_\oplus B_0^{-1} P_\oplus^T\right]^{-1} H^T\right\}}{Tr\left\{\delta y^T \delta y\right\} - Tr\left\{R\right\} - Tr\left\{HB_m H^T\right\}/\alpha} \quad (30)$$

を用いて求める．(30) の分子など Tr の計算量が多い項は，モンテカルロ法によって近似計算する（Bai *et al*. 1996 [3]）．

背景誤差共分散の融合表現（25 式，あるいは，26 式）は，北米大陸西岸のモントレー湾におけるグライダーデータの同化実験に用いられた [4]．この実験では，異なる時刻からの予測値に基づく 30 個のアンサンブルメンバーを逐次的に生成しながら 3 次元変分法によるデータ同化が行われた．その結果，1–2 個の統計的に有意な固有モードがアンサンブルから検出され，背景誤差共分散の融合表現によって 15–20% のスキル改善がみられた．推定された β/α は 1–2 程度と，アンサンブルによる背景誤差共分散が持つ重みは経験的背景誤差共分散のそれに比べて相対的に大きい結果となった．

私たちは，アンサンブルカルマンフィルターを日本南岸域の黒潮流路変動に適用したが，生成したアンサンブルメンバーの海面高度場に対し，(24) と同様な共分散の近似を行った．上位の固有モードのうち，大蛇行と非大蛇行流路に対応するモードを抽出することができた．またこれらのモードは，混合分布モデル（Bishop 2006 [7]）で分類された流路場とも整合的であった（Aoki *et al*. 2020a）．

背景誤差共分散の融合表現は，アンサンブルカルマンフィルターと四次元アンサンブル変分法の融合にも使えそうである．通常の四次元アンサンブル変分法では経験的な背景誤差共分散のみを用いている．四次元アンサンブル変分法で生成した解析値に，アンサンブルカルマンフィルターによって生成したアンサンブル摂動を加えてアンサンブル予測を行い，次の解析の際の背景誤差に融合させるのである．私たちの経験では，日本南岸域の黒潮流路を対象としたアンサンブル四次元変分法の適用において 9 日間の時間ウインドウの初期値背景誤差による感度は十分に大きく（Miyazawa *et al*. 2020），背景誤差の融合表現によるスキル改善の可能性があると考えられる．アンサンブルによる背景誤差共分散の表現は，等方的かつ静的な経験的表現に対し,「流れに依存する」動的な誤差の情報を与えるという特徴がある．

4. アンサンブルによる観測データ寄与評価

多くの場合アンサンブルは，前の解析時刻から観測データ同化によって生成された初期値からの摂動によって生成されるので，その品質は観測

データの種類や分布に影響されるはずである．今後，アンサンブルによる背景誤差共分散と，観測データの種類や分布との関連をより詳しく調べていきたいと考えている．

アンサンブルは，モデル予測に対する個々の観測データの寄与を評価することに用いることもできる（Kalnay *et al*. 2012 [8]）．ある時刻（t=0）に同化した観測データの寄与を評価するため，同化した結果を初期値とする予測結果と t=0 での解析値の差 $e_{t|0} = x_{t|0}^{f} - x_{t|0}^{a}$ と，それ以前（D 日前）からの予測で t=0 に観測データを同化しないまま継続した予測結果と t=0 での解析値の差 $e_{t|-D} = x_{t|-D}^{f} - x_{t|-D}^{a}$ を用いて，以下で定義する量（二乗ノルムの差）を計算する．

$$J = \left[e_{t|0}^{T} e_{t|0} - e_{t|-D}^{T} e_{t|-D} \right] \tag{31}$$

これは，

$$\begin{aligned} J &= \left[e_{t|0} - e_{t|-D} \right]^{T} \left[e_{t|0} + e_{t|-D} \right] \\ &= \left[x_{t|0}^{f} - x_{t|-D}^{f} \right]^{T} \left[e_{t|0} + e_{t|-D} \right] \\ &\sim \left[M\left(x_{t|0}^{a} - x_{t|0}^{f} \right) \right]^{T} \left[e_{t|0} + e_{t|-D} \right] \\ &= \left[MK\delta y_{t|0} \right]^{T} \left[e_{t|0} + e_{t|-D} \right] \\ &= \delta y_{t|0}{}^{T} K^{T} M^{T} \left[e_{t|0} + e_{t|-D} \right] \end{aligned} \tag{32}$$

と表わされ（K はカルマンゲイン），通常はアジョイントモデル M^{T} を用いて算出される（Forecast Sensitivity to Observation（FSO））．これに対し Kalnay *et al*.（2012）[8] は，

$MK = \dfrac{1}{m-1} MX_{t|0}^{a} X_{0}^{aT} H^{T} R^{-1} \sim X_{t|0}^{f} Y_{t|0}^{aT} R^{-1}$ を用いて

$$J \sim \frac{1}{m-1} \delta y_{t|0}^{T} R^{-1} Y_{t|0}^{a} X_{t|0}^{fT} \left[e_{t|0} + e_{t|-D} \right] \tag{33}$$

とする近似を提案した（Ensemble Forecast Sensitivity to Observation（EFSO））．(33) は，解析時（t=0）の各観測データの寄与に分解できるので，線形の時間発展が仮定できる期間内においては各観測データの予測に与える寄与を定量的に評価できる．また，(33) はカルマンゲインを明示的に含んでいないので，形式的には様々なアンサンブル同化法に適用できるという利便性がある．ただし，(33) の算出のため，アンサンブル予測結果 $X_{t|0}^{f} = (\delta x_{t|0,i=1}^{f}, \delta x_{t|0,i=2}^{f}, ..., \delta x_{t|0,i=m}^{f})$ を予測期間内においてすべて保存しておく必要がある．最近は，FSO (32) も EFSO (33) もともに気象・海洋予測の分野でよく使われるようになっている（例えば，気象では Kotsuki *et al*.（2019）[9]，海洋では Powell（2017）[10] など）．Kim and Kim（2019）[11] は，韓国気象庁の現業モデルの四次元変分法同化において FSO を適用し，様々な観測データが全球スケールの気象予測にもたらす影響を評価した．興味深いことは，アンサンブルによる背景誤差共分散の融合表現（第 3 節）を導入すると，ほぼ全種類の観測データの寄与が向上する（Fig. 2 [11] 参照）が，その向上度合いは観測データの種類によって相当程度異なることである．観測データの寄与はモデルと同化手法の選択に強く依存するので，その解釈に注意が必要である．一般に，モデルと同化手法が，予測の対象となる現象の種類や時空間スケールに応じて選択されることを考えると，それぞれの観測データの寄与度は結局，予測の対象となる現象によって大きく違ってくる可能性があることに注意したい．ある観測データの有効性を評価する場合，決してひとつのモデルや同化手法のみで判断してはならないということである．

5. まとめ

現在，私たちはモデルのアンサンブルを効率的に生成するためのモデル高速化や計算環境の整備を進めている．アンサンブルによる予測が可能になれば，予測の当たりはずれについて定量的な診断を行うことができる．さらに，上述したようなアンサンブルの活用によって観測との協働を積極的に進めていきたい．また，アンサンブルによっ

てモデルシミュレーション結果の確率分布を精度よく表現できるようになれば，確率分布を用いたレイノルズ応力の表現（Aoki *et al*. 2020b，改訂中）や，エネルギー解析へのアンサンブルの活用（Aoki *et al*. 2020a）などシミュレーション結果の様々な解析にも応用できると思われる．本稿では触れなかったが，より本質的な事柄として，品質の良い（未知の確率分布をより正確に表現する）アンサンブルをどのように生成していくかという課題があるが，これについても今後追求していきたい．

参考文献

[1] 伊藤, 身延 2010: 気象学と海洋物理学で用いられるデータ解析法, 気象研究ノート, 221.
[2] Fujii *et al*. 2008: Journal of Geophysical Research, 113, C07026.
[3] Enomoto *et al*. 2015: Journal of Meteorological Society of Japan, 93, 199-213.
[4] Yaremchuk *et al*. 2011: Monthly Weather Review, 139, 1879-1890.
[5] Evensen and van Leeuwem 1996: Monthly Weather Review, 124, 85-96.
[6] Bai *et al*. 1996: Journal of Computational and Applied Mathematics, 74, 71-89.
[7] Bishop, 2006: Pattern recognition and machine learning (information science and statistics), Springer-Verlag.
[8] Kalnay *et al*. 2012: Tellus A Dynamic Meteorology and Oceanography, 64, 18462.
[9] Kotsuki *et al*. 2019: SOLA, 15A, 1-7.
[10] Powell 2017: Journal of Geophysical Research Oceans, 122, 8427-8444.
[11] Kim and Kim 2019: Journal of Atmospheric and Oceanic Technology, 36, 1563-1575.