

データ同化による数値モデルと観測データの統融合

講演者: 石川洋一 (海洋研究開発機構・地球情報研究センター)

報告者: 野口峻佑 (地球物理学教室 M2)

1 講演概要

「データ同化」とは、観測データと数値モデルを組み合わせ、現実的なデータセットを作るための手法である。本講演では、まずデータ同化の背景について、丁寧な導入がなされた後に、その手法の数学的基礎が概観され、実際に用いられているデータ同化手法の解説がなされた。最後に、データ同化の応用として、講演者らのグループが行っている海洋循環変動・気候変動研究が紹介された。

2 内容

2.1 データ同化の背景

講演では、まず導入として、データ同化が必要となる背景とその目的についての説明がなされた。

データ同化とは、観測データと数値モデルを組み合わせ、現実的なデータセットを作成する手法のことである(図1)。データ同化の目的を大きく4つに分けると以下のようになる。

1. 均一な統合時系列データセットの作成
2. 数値予報の精度向上のための初期値・境界値作成
3. 数値モデルで用いているパラメータの最適化
4. 観測システムの評価・改善のための情報提供

データ同化は、数値予報に必要な初期値の正確な推定を目的に、その研究が始まった。数値予報は、運動方程式やエネルギー保存則などの支配方程式を離散化した数値モデルにより行われるが、その実行のためには初期値・境界値が必要である。初期値としては観測値が用いられるが、得られる観測データには誤差が含まれ、また、必ずしも数値モデルの計算格子にあるわけではない。そこで、以前の時刻からの予報値と得られた観測値を組み合わせ、数値モデルと整合的なデータセットを作成する。

この操作は、観測データ側からみると、数値モデルを利用した計算格子上への(4次元的な)内外挿、異なる種類の観測データの統合をしたものとみなせ、また数値モデル側からみると、モデルの入力パラメータ(初期値・境界値など)のチューニングの自動化をしたものとみなせる。

数値天気予報が発達した気象の分野では2に重点をおいた研究が、十分な観測を得ることが難しい海洋の分野では1に重点をおいた研究がなされてきた。

2.2 データ同化の基礎

次に、データ同化手法の数学的基礎となる、線形不偏最適推定 (Best Linear Unbiased Estimation, BLUE) 理論について、そのエッセンスが述べられ、その多変量・多次元への拡張である、実際に用いられるデータ同化手法についての解説がなされた。

データ同化には、大きく分けて、Sequential な手法と Iterative な手法がある(図2)。講演では主に、前者の代表としてカルマンフィルター、後者の代表として4次元変分法が紹介された。

カルマンフィルターでは、誤差についての行列の時間発展を予報モデルを用いて求めるため、流れ場の物理的状態に即した誤差を考えることができる。しかし、大気・海洋モデルのような自由度の大きなモデルでは、この行列の計算負荷が大きすぎる。このため、複数個の予報を走らせるアンサンブル予報を実行し、その複数個の予報によって行列を近似するアンサンブルカルマンフィルターが主に用いられる。

4次元変分法では、随伴(アジョイント)モデルと呼ばれるモデルを用いて、誤差情報を時間方向逆向きに伝播させ、同化期間における修正を繰り返すことにより最適解を得る。

予報モデルが線形な場合には、カルマンフィルターによる推定値と4次元変分法による推定値は一致する。しかし、非線形な場合には、両者が一致する保証はなく、アンサンブルカルマンフィルターと4次元変分法の比較は、ホットな研究トピックである。

2.3 データ同化の応用

最後に、データ同化の応用として、講演者らのグループが行っている海洋循環変動・気候変動研究が4つ紹介された。ちなみにこれらの紹介された研究は、データ同化の目的で挙げた4つの項目に対応している。

まず、均一な統合時系列データセットの作成の例として、海面水温・亜表層水温場の同化結果が示された。沖合いにおける観測はまばらであるが、数値モデルによる計算結果と統合され、観測でみられた特徴を再現しつつも欠損のない均質なデータが作成されている(図3)。

次に、数値予報の精度向上のための初期値・境界値作成の例として、大気海洋結合モデルによるエルニーニョ・北

太平洋十年振動の予測結果が示された。予報開始3ヶ月前のデータを同化することにより、予報成績が改善されており、約15ヶ月前からエルニーニョの予測に成功している(図4)。

また、数値モデルで用いているパラメータの最適化の例として、海洋の生態系モデルのパラメータの最適化についても触れた(図5)。数値モデル中では、数多のプランクトンの生態系としての複雑な振る舞いが、1000個近くのパラメータを用いて記述されており、その海域毎のパラメータの最適化の結果が紹介された。

観測システムの評価・改善のための情報提供の例としては、Argo フロートによる観測データのインパクト評価が紹介された。Argo フロートによる観測データのあり/なしの結果が示され、新たな観測データを取り入れることの影響が示された(図6)。

2.4 データ同化の今後

気象・海洋の分野で発展してきたデータ同化であるが、最近では地震などの固体地球分野でもデータ同化システムの開発が行われており、地球科学において一般的なものとなってきている。今後も、計算機性能の向上や数値モデルの発展、新しい観測結果の導入とともに、新たな分野への応用の機会が増えていくことが期待される。

3 報告者感想

現在、気象におけるコミュニティモデルの1つであるWRFでは、データ同化ツールWRFDAが標準装備されるなど、データ同化のモデルへの標準パッケージ化が進んでいる。このことにより、今後、データ同化を応用した研究が大きく進展していくことが期待されるが、同時に多くのユーザーにとっては、データ同化がブラックボックスとなってしまうことが危惧される。データ同化の研究自体、強非線形性への対応やアンサンブル手法による確率的予報など、魅力的な課題が数多く残っている。リバースエンジニアリングの難易度を下げるなどの、ユーザー側からメーカー側への「橋渡し」への配慮が重要であるように感じられた。

* 図は講演者のスライドより引用させてもらった。

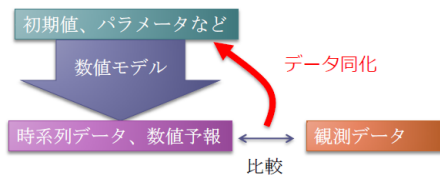


図1: データ同化の概念図。

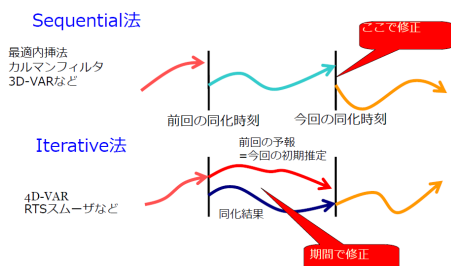


図2: データ同化手法の分類。

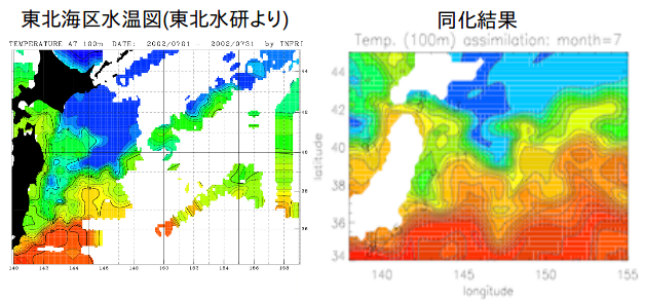


図3: 水深100mにおける7月の水温場の図。左図が観測値、右側が同化結果。

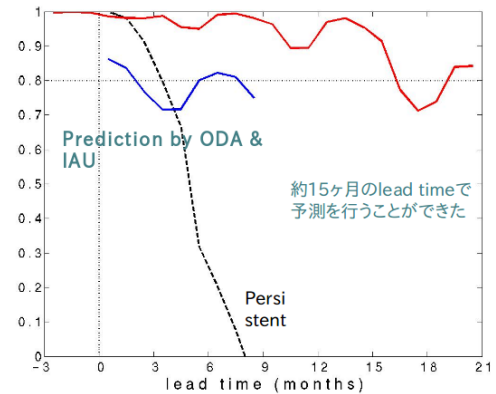


図4: アノマリー相関(パターンの一致度の指標。値が1に近いほど良)の時系列。赤線が予報開始3ヶ月前から大気海洋合同化システムにより同化した結果、青線が海洋データ同化システムによる結果。

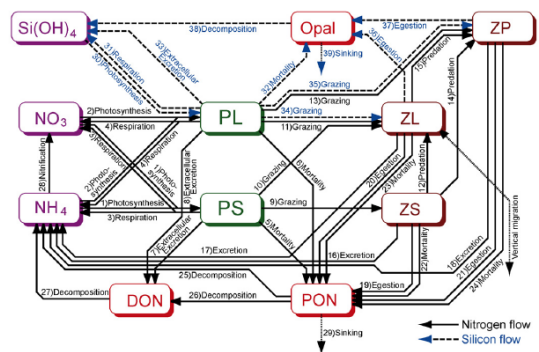


図5: 海洋生態系モデルのフローチャート。

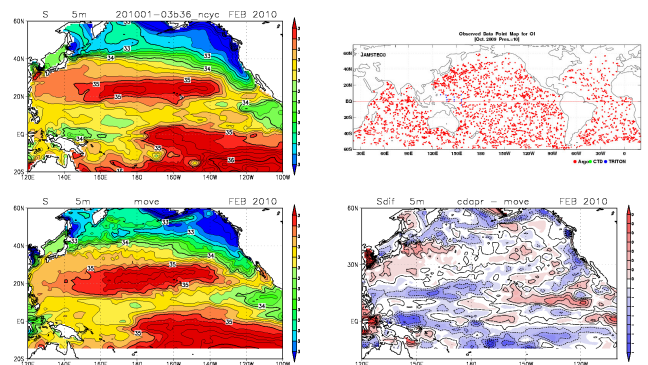


図6: Argo フロートによる観測データのインパクト評価。海面塩分の結果。左上: Argo データなし。左下: Argo データあり(MOVE_NP)。右上: Argo データ観測地点(赤丸)。右下: Argo データなしの MOVE_NP との差。