

4. データ同化への誘い

三好 建正*

1. はじめに

気象学において、データ同化研究が近年盛んになっている。そもそも気象学は現象学であって、実際に自然大気で起こっている現象を観測することで、気象に関する科学的知見を深めてきた。新たな衛星などの観測測器を開発・運用したり、これまで観測が少なかった場所へ行って観測したりするなど、実際に自然で起こる新たな現象を観測したいという観測研究者は多く、気象学の本流を継承している。一方で、数値予報モデル（以下単に「モデル」）を使った研究手法も近年の気象学の本流を成すといえよう。ある既知の大気現象が、モデルでは表現されたりされなかったりする。この原因を考えることで、そのメカニズム解明に役立つ。また、自然では起こりえない設定、たとえばコリオリ因子を0にしたり、地形を変えたり、そういう実験を行うことで、自然に存在するコリオリ力や地形がどういう役割を果たしているか知ることができる。また、大気中の二酸化炭素濃度を自由に設定して模擬実験することで、気候変動の予測にも応用できる。

以上のように、気象学には、観測およびモデルという大きく2つの方法論がある。では、データ同化とは何か。データ同化は、観測とモデルの双方を使って、モデルで表現される大気状態が現実をよく表すようにする。観測データをモデルの状態変数に同化することから、データ同化と呼ばれる。データ同化は、気象学の2つの方法論を橋渡しするものと見ることができ、双方の相乗効果を生むことで気象学の発展に貢献するものと期待できる。また、データ同化は後に述

べるように統計数理的手法であり、伝統的な気象学的アプローチとは異なった視点を与えうるといえる観点でも、気象学の幅を広げるものと期待できる。

このように、データ同化は異分野間の橋渡しをしたり、新たな視点を与えたりするものだから、幅広い研究者に知っていただき、データ同化が各分野でどう役立つのか考えていただくことは、有益だろう。また、様々な背景を持つ研究者がこのような観点で互いに議論することも、とても重要である。本稿は、このような議論を行うにあたって、共通の土台を与えることを目的としている。

本稿は、著者の知識の偏りが原因で網羅的でなく、数値天気予報に偏った解説となっていることをお詫びしたい。データ同化は、気象以外にも、海洋や陸面を含めた地球科学に留まらず、火星などの惑星大気、太陽放射に起因する宇宙天気など、様々な分野でもこれまで応用され、またさらに新しい分野にも応用の幅が広がってきている。本稿が数値天気予報に偏っているのは、単に著者の無知によるものであり、議論の対象を狭い分野に限ることを意図するものではないことをご理解いただければ幸いである。

2. データ同化のプロダクト

ここではまず、データ同化の具体的なイメージをつかむため、データ同化が何を作り出すかを見たい。

データ同化は、観測とモデルの双方を最大限に生かし、現実大気の推定値を作り出す。データ同化が作り出すプロダクトで、気象学研究者が最も馴染み深いものは、再解析データであろう。再解析データとは、過去数十年に渡る大気状態の推定値データベースである。大気状態は、地球大気を格子点に分割して、各格子点上の各変数（風速成分、気温、湿度等）の値の集合として表現される。再解析では、通常6時間毎のス

* メリーランド大学大気海洋科学部。

miyoshi@atmos.umd.edu

© 2010 日本気象学会

ナップショットとしてデータが生成されており、日本時間でいうと午前・午後の3時と9時の毎日の大気状態がわかることになる。過去数十年にわたる任意の時刻の任意の場所で大気状態の推定量が得られることは、様々な研究にとって非常に有用である。NCEP/NCAR 再解析 (National Centers for Environmental Prediction / National Center for Atmospheric Research reanalysis, Kalnay *et al.* 1996) は、NCEPの現業数値天気予報システムに基づいて世界で初めて作成された再解析データで、数多くの研究で使われてきた。Kalnay *et al.* (1996) による論文が地球科学で最も引用されている論文の1つであることが、再解析の有用性、ひいてはデータ同化の重要性を証明している。

再解析データはデータ同化プロダクトの一例であり、他にも、現業数値天気予報における客観解析データという実用上非常に重要なデータ同化プロダクトがある。現業数値天気予報における客観解析データは、モデルの初期値となる。モデルの性能が向上するにつれ、初期値が予報精度に与える影響が相対的に大きくなってきた。近年の数値天気予報精度の向上において、データ同化に関する開発は重要な役割を果たしている。このことが、データ同化研究が近年盛んである主な理由となっている。

3. データ同化のプロセス

再解析も現業客観解析も、基本的には同じデータ同化手法を用いて作られる。時空間的にくまなく分布したモデル格子点値に比べて観測は一般にまばらすぎる。このため、観測値だけを使ってモデル格子点上に配置されたすべての変数を推定することは、非常に困難である。そこで、直近のモデル予報値 (全球大気の場合は通常6時間予報値) と観測値とを組み合わせる。予報値と観測値は、現実大気に関して少なくともある程度は独立した情報を持つから、両者を組み合わせることで、より情報量の多い解析値を作り出すことができる。この統計数理的プロセスがデータ同化である。

データ同化では、観測値や予報値は、確率的に分布した量 (確率変数) とみなされる。確率変数は確率密度関数というより高次元な量によって表現されるため、我々がこれを直接認識することは難しい。通常我々が認識する観測値や予報値は、確率変数の期待値とみなされる。

確率変数の考え方は、データ同化のプロセスで非常に重要である。その理解を助けるため、簡単な例を示したい。まず、観測を1つだけ行うことを考える。観測値は1つだけ得られるわけだが、データ同化ではこれを1次元の確率変数に関する1回の試行とみなす。試行を無限回繰り返せば、1次元的に分布した確率密度関数を正確に推定できる。観測の誤差がガウス分布に従うと仮定すれば、その期待値と標準偏差の2つのパラメータのみによって、この1次元の関数を表現できる。次に2つの観測を行うことを考える。これらの観測が互いに独立なら、それぞれの観測に対して独立に確率密度関数が存在することになる。一般には、2次元の確率密度関数を考えればよい。ガウス分布を仮定すると、2つの観測それぞれの期待値と、2行2列の共分散行列で一意に定まる。観測が独立な場合、共分散行列の非対角成分は0である。言い換えれば、非対角成分を考慮することで、互いに独立でない複数の観測も考慮に入れることができる。同様に、 n 個の観測を行った場合は、 n 次元の確率密度関数を考えればよく、ガウス分布を仮定すれば、 n 個の期待値と $n \times n$ の共分散行列を決めればよい。予報値が n 次元の確率変数とすれば、 n 個の観測を行った場合と同様に考えればよい。

さて、データ同化において観測値と予報値を組み合わせる際、解析値を得る際には、解析値の確率密度関数を、観測値と予報値の確率密度関数の積によって与える。これは、数学的にはベイズの定理として知られており、ベイズ推定と呼ばれる (詳しくは例えば、淡路ほか 2009を参照)。2つのガウス分布関数の積はガウス分布関数であり、この分散は、元の2つの分散のどちらよりも小さい。つまり、ガウス分布を仮定すれば、観測値と予報値を組み合わせることで得られる解析値は、そのどちらよりも確からしいことになる。

実際の大気モデルの自由度は1000万もの桁になるため、予報値の確率密度関数を直接扱うことは非現実的である。ガウス分布を仮定しても、共分散行列の成分数はモデル自由度の2乗で増えるため、大気特有の特徴を生かすなどして共分散行列の性質を扱いやすくして計算効率を上げる工夫がなされている。具体的には、共分散行列が時空間的に一様だと仮定したり、共分散行列の成分間に空間的なある構造を仮定したり、風と質量の間にあるバランス関係 (地衡流バランス) などの条件を課したりして、大幅な効率化を行う。

データ同化手法の研究として近年盛んなのは、上で

述べた様々な効率化の仮定のうち時空間一様性についての束縛を緩和し、日々変動する大気状態に応じた確率密度関数を考慮することである。例えば、アンサンブル・カルマンフィルタ (EnKF, より詳しくは例えば三好 2005を参照) は、アンサンブル予報を使って日々変動する予報誤差共分散を推定し、データ同化を行う。このような新しい手法を研究するのは、これまでの手法が静的であったり非効率であったりする点を改良し、解析精度および計算効率を向上する余地がまだ多くあるからである。

4. データ同化応用研究の広がり

冒頭で述べたように、データ同化は観測とモデルをつなぐ橋のようなものである。モデル研究者にとっては、モデルに観測を同化することで、数値天気予報のようにそのモデルに適した初期値が得られ、よりよい現実大気のシミュレーションができるようになる。そのモデルによる現実大気の推定値が得られること自体も魅力で、他の再解析データなどと比較することで、そのモデル特有の誤差バイアスなどの情報が得られる。こういった情報はモデルの改善に役立つとともに、より現実大気にフィットするようなモデリング研究を通じて大気現象のメカニズムの解明にもつながる。

一方、観測研究者にとっても、データ同化の結果である解析データは有益である。データ同化によって観測に最もよくフィットするように作成された時空間的に密な解析データは、まばらな観測データを強力に補う。また、モデル変数の関数として表されるどんな物理量の観測データも扱える。たとえば、衛星から観測されるマイクロ波の放射輝度といった物理量を観測した場合でも、それを最適に取り込んだ解析データがモデル格子点値として得られる。

また、新たな観測システムをデザインする際にもデータ同化は有効である。これには、データ同化手法を研究する際に用いられる実験手法を応用する。架空の観測システムを想定して行う「観測システムシミュレーション実験 (OSSE: Observing Systems Simulation Experiment, 例えば Atlas *et al.* 1985)」と呼ばれる手法である。これにより、例えば新しい地球観測衛星を打ち上げる前に、その衛星の導入により解析や予報に対してどれだけの効果が見込まれるか等、コストのかかる観測の計画段階で、そのインパクトを事前に定量的に推定することが可能となる。

このように、データ同化が間に入ることによって、観測研究とモデル研究は、両者の利点を生かしつつ、互いにより明確に意識できるようになる。喩えていうなら、データ同化の橋は、観測からモデルへの一方通行ではなく、双方向なものである。

数値天気予報にデータ同化は不可欠だから、多くの数値天気予報モデル研究者にとって、データ同化は比較的馴染み深いものといえる。最近はこの枠にとらわれず、大気微量成分の予測モデル等 (Lahoz *et al.* 2007)、様々なモデルを用いたデータ同化研究も盛んになってきている。また、データ同化を使った観測研究者の取り組みも既に始まっている。たとえば、Moteki *et al.* (2007) は、熱帯西太平洋での観測プロジェクト PALAU2005の観測データを同化し、これを同化しない場合との違いを見て、観測のインパクト研究を行った。このような観測インパクト調査は、今後の観測研究を計画するに当たり、有益である。また、Inoue *et al.* (2009) は、北極海の海水上にブイ観測点が設置されていることに着目し、温暖化による北極海の海水減少によるブイ観測点の減少が大気客観解析に与えるインパクトについて論じている。これら実際の観測システムのインパクトを調べる実験は、観測システム実験 (OSE: Observing Systems Experiment) として知られている。

5. 議論の始まりにあたって

データ同化手法の研究は、それ自体としては統計数理的な側面が強く、多くの大気・海洋・水文分野等の研究者にとって取っ付きにくいと思われてきたかもしれない。しかし、その成果を応用することで、多くの研究者にとって有益な研究ができそうだ。さらに、データ同化を橋とした観測研究とモデル研究との間の交流は、新たな相乗効果を生み出すものと期待できる。様々な分野の研究者にとって、データ同化をどのように応用できるのか、今後ますます活発な意見や情報の交換が期待されるところである。

参考文献

- Atlas, R., E. Kalnay, W. E. Baker, J. Susskind, D. Reuter and M. Halem, 1985: Simulation studies of the impact of future observing systems on weather prediction. Proc. Seventh Conf. on NWP, 145-151.
- 淡路敏之, 蒲池政文, 池田元美, 石川洋一, 2009: データ同化~観測・実験とモデルを融合するイノベーション

- ～. 京都大学学術出版会, 284pp.
- Inoue, J., T. Enomoto, T. Miyoshi and S. Yamane, 2009: Impact of observations from Arctic drifting buoys on the reanalysis of surface fields. *Geophys. Res. Lett.*, **36**, L08501, doi : 10.1029/2009GL037380.
- Kalnay, E. *et al.*, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteor. Soc.*, **77**, 437-471.
- Lahoz, W. A., Q. Errera, R. Swinbank and D. Fonteyn, 2007: Data assimilation of stratospheric constituents: A review. *Atm. Chem. Phys.*, **7**, 5745-5773.
- 三好建正, 2005: アンサンブル・カルマンフィルタ—データ同化とアンサンブル予報の接点—. *天気*, **52**, 93-104.
- Moteki, Q. *et al.*, 2007: The impact of the assimilation of dropsonde observations during PALAU2005 in ALERA. *SOLA*, **3**, 97-100.
-