

# 知っておきたいキーワード

## データ同化

越後 富夫<sup>†</sup><sup>†</sup> 大阪電気通信大学

"Data Assimilation" by Tomio Echigo (Osaka Electro-Communication University, Osaka)

キーワード：気象学，4次元変分法，アンサンブル・カルマンフィルタ，深層学習，画像インペインティング

### まえがき

近年、コンピュータの演算能力の向上により、膨大なデータを解釈し、帰納的に事象を読み解くアプローチであるデータサイエンスが注目されています。一方で、複雑な現象の予測、解析において、実際の計測が容易でない場合も多く、そのような場合には、現象を数学的に表現した数理モデルを用い

たシミュレーションが非常に重要と なっています。

機械学習は、コンピュータに膨大なデータを読み込ませることで現象のパターンや特徴を抽出する技術であり、近未来予測やシステム制御、異常検知に活用されています。その過程では、大量のデータセットが入力可能である場合、特定のアルゴリズムを利用して数理モデルを正確に構築することがで

き、正確な予測分析が行われます。しかしながら、日常的に発生頻度が低いイベント、例えば、飛行機事故や巨大津波などの大規模災害では、データが集まりにくく、数理モデルにおける正確なパラメータを決定することができないため、予測性能が不十分であることが問題になります。この課題を解決するのがデータ同化です。

### データ同化

データ同化は、代表として気象学、海洋学の分野で活発に用いられている手法です。地球システムの再現、解明、予測には、物理法則に基づいた時空間シミュレーションモデルを構成することが重要です。気象現象では、水蒸気量や風向、風速に加え、地表の形状などのさまざまな要因が互いに影響することで、内部状態を表現する数多くのパラメータは非線形性を伴い、正確なパラメータ推定は簡単には行えません。さらにシミュレーションではすべての要因が含まれているわけではないため、実際の観測値とシミュレーション推定値が一致することは稀です。そこで、図1に示すように、

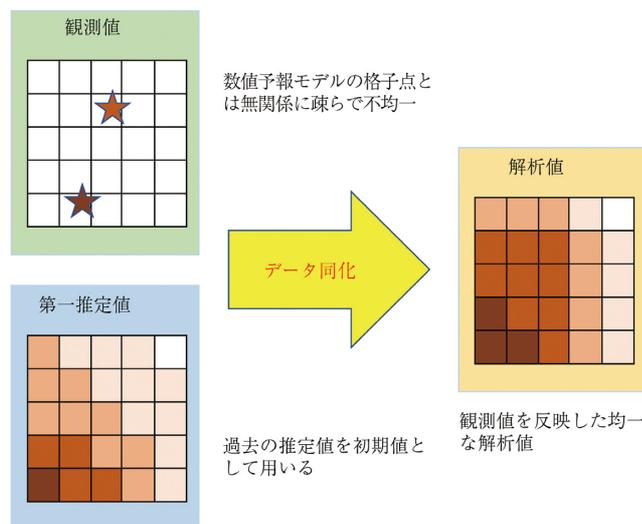


図1 データ同化の模式図

☞ 数値予報モデルの格子点と比較して疎らでかつ分布も不均一な観測値と、数値予報モデルに利用可能な分解能を持つ過去のシミュレーション結果を第一推定値として用い、データ同化することで、ある時刻の物理量分布で

ある解析値を求めます。ただし、データ同化の際には、予報値だけでなく、観測値にも誤差があると考えられるため、単純に観測値を解析値に置き換えるのではなく、観測誤差と推定誤差をブレンドすることで解析値を求めま

す。その結果、データ同化によって、観測値の得られない時間、場所における物理量をもっともらしく推定することができ、時間的、空間的に均質なデータ生成ができます。

### データ同化手法

#### (1) 非逐次型手法

以前から主流として用いられてきた手法の一つが「最適内挿法」です。この手法は第一推定値から観測値のずれを観測誤差と予報誤差の統計情報を基に第一推定値の格子に修正して内挿する方法です。気象現象では、最近では大量の衛星観測データを利用することが多いため、観測データを直接扱えない最適内挿法は大気データのデータ同化に用いられなくなっています。次に、大気現象は3次元空間に存在することから、さまざまな時刻における観測データを基に、評価関数を用いて繰り返し計算により解析値を求める方法が「3次元変分法」です。数値予報モデルを実行

しないので、計算量が少ないですが、大気場の力学的な時間変化が考慮されないため、力学的にバランスの取れた解析値が得られない欠点があります。一方、大気現象は3次元空間において時々刻々と変化しています。空間のさまざまな地点、任意の時刻において観測されるデータを空間+時間の4次元で扱うことで、より精度の高い解析値を得るのが「4次元変分法」です<sup>1)</sup>。4次元変分法では、数値予報モデルの時間積分値がデータ同化期間中の観測値に最も合うように初期値を修正していきます。繰り返し処理で第一推定値を少しずつ修正し、評価関数を用いて、第一推定値と観測値の両方に最も近い数値予報モデルの結果を最適な解析値と定めます。

#### (2) 逐次型手法

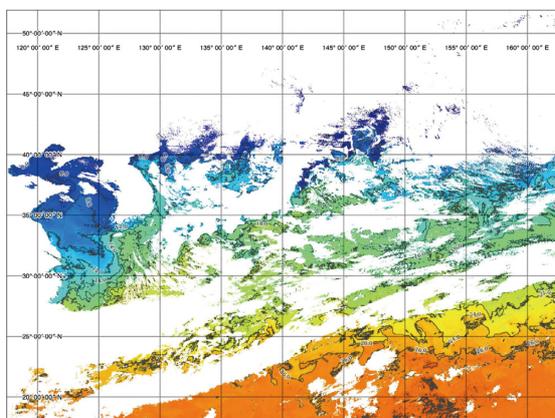
逐次型手法では、観測値が得られるたびに、その時刻での状態変数を推定します。そのためにカルマンフィルタが使われましたが、カルマンフィルタはモデルが線形で誤差がガウス分布する場合には最適になりますが、そうでない場合には問題になります。そこで少数のサンプルで誤差を表現するアンサンブル予報を利用した「アンサンブル・カルマンフィルタ」が提案されました<sup>2)</sup>。この手法は、アンサンブル予報のメンバから予報誤差を見積もり、解析値の誤差が最小となるように最適解を求める手法です。実際に、推定する状態変数ベクトルの次元が非常に大きい問題に適用しても有効であったことが報告されています<sup>3)</sup>。

### 深層学習を利用したデータ同化

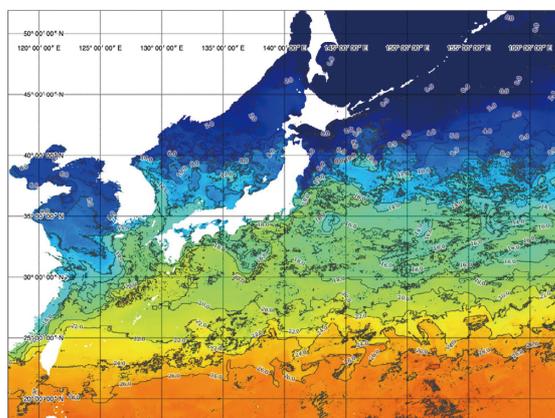
海洋環境の計測において、気象衛星は広域をカバーできることで有効です。図2(a)は、気象衛星で観測された情報

を基にした海表面の水温画像です。ここでは、近赤外線の反射から水温を推定していますが、雲がかかっている領域は水温計測ができていません。この領域に対し、物理シミュレーションを

利用して欠損した水温情報を推定する手法が提案されていますが、スーパーコンピュータによる多大な計算コストが必要になっています。そこで、この画像を欠損画像として、☞



(a) 雲によって部分的に隠された海洋画像



(b) 修復した海洋画像

図2 画像インペインティングによる欠損海洋画像の修復  
(出典：飯山将見：映像情報メディア学会誌，73，5，pp.946-949（2019））

☞ 画像を修復する画像インペインティング手法を適用します。この技術は従来から多くの研究がありましたが、近年、深層学習を利用した手法が良い性能を示しています。深層学習では教師データとして大量の画像を必要

としますが、海水温画像で雲一つない画像を用意することが難しいため、欠損画像も教師データに用いることができるように、欠損領域をマスクとして入力するネットワークの改良を行っています。また、衛星画像にはノイズが

多く、ノイズを除去して復元するために、物理シミュレーションによって生成する水温データを活用した敵対的生成ネットワークを用いています。復元処理を衛星画像全体に繰り返し適用することで復元した画像が図2(b)です<sup>4)</sup>。

### その他の応用

データ同化は、先に記述したように気象学、海洋学で用いられましたが、もっと身近な場所でも有望視されています。その一つが快適なオフィス空間を提供するスマートオフィスです。スマートオフィスでは、温度管理が重要で、人が行動する空間の温度を測定し、それを基に空調設備を制御することが必要になります。現状の空調設備は、設定温度は空調機の吸込口の温度を基にしていますが、空間の温度を測定するには、離れた場所からの測定は不可能で、必要とする位置、高さに温度センサを設置する必要があります。そのような位置は人の行動の障害になることから、できるだけ少数の温度センサで空間の温度を推定する必要があります。そこで、部屋の数箇所に温度センサを設置し、センサのない地点の温度をデータ同化によって推定することにします。図3(a)に示すように、部屋を2次元グリッドに分割し、設置した温度センサがあるグリッドを画素とみ

なすことができ、部屋の温度分布を画像として扱えます。近年、粗い解像度の動画を、深層学習を用いて高精細化する手法 (tecogan)<sup>5)</sup> が提案されており、この手法を用いて温度分布の2次元高精細動画を生成した結果が図3(b)です。一方、温度分布を検出する方法にサーモグラフィがありますが、サーモグラフィが検出する温度は物体表面の輻射熱であり、部屋の床、壁、天井、什器などの構造物の表面温度になるため、空間の温度は測定できません。温

度センサから構造物までの距離はそれぞれ異なり、構造物からの輻射熱が空気を伝わり温度センサに到達する関係性をニューラルネットで表現します。両者間の変換パラメータを学習によって求めることで、サーモグラフィを入力として与えると、元の温度センサが存在していた位置の温度データが推定できます。その結果、tecoganを使うと2次元高精細の熱分布動画が得られます。

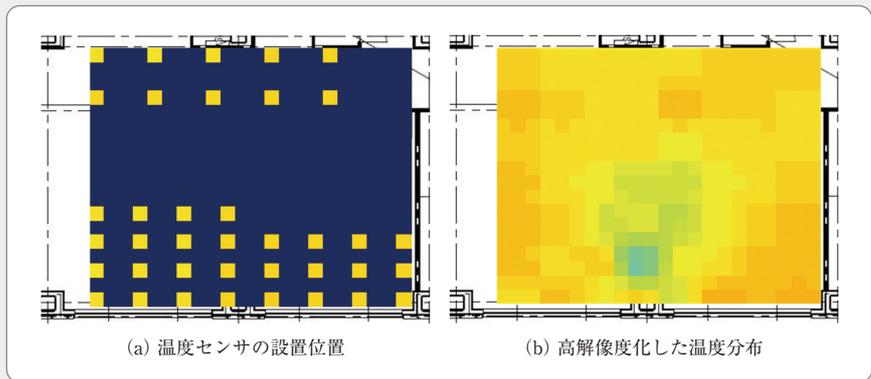


図3 温度センサを基にした高精細度動画生成

### むすび

ICTの発展に伴い、大量のデータが獲得できるようになりましたが、データがカバーする領域は必ずしも均一ではありません。欠損している領域のデータも過去および周辺のデータから推定できる可能性があります。それを実現するデータ同化は、今後ますます応用が広がり、用途に適した手法が提案されると思います。

(2020年3月31日受付)

### 参考文献

- 1) 露木義：数値予報課報告別冊，48，気象庁予報部，1-16 (2002)
- 2) G.Evensen: "The ensemble Kalman filter: Theoretical formulation and practical implementation", *Ocean Dynamics*, 53, 343-367 (2003)
- 3) K.A. Lisæter, J. Rosanova and G. Evensen: "Assimilation of ice concentration in a coupled ice-ocean model, using the ensemble Kalman filter", *Ocean Dynamica*, 53, 368-388 (2003)
- 4) 飯山将見：“水産業における画像・映像処理技術の応用”，*映像学誌*，9 (2019)
- 5) M. Chu, Y. Xie, L. Leal-Taixe and N. Thuerey: "Temporally coherent gans for video super-resolution (tecogan)", *arXiv preprint arXiv:1811.09393* (2018)



### 越後 富夫

1982年、日本アイ・ビー・エム(株)入社。東京基礎研究所にて、ロボットビジョン、映像メディア処理の研究に従事。2003年、大阪大学大学院基礎工学

研究科博士後期課程修了。同年、同大学産業科学研究所客員教授。医用画像診断支援の研究に従事。2006年より、大阪電気通信大学教授。2017年、大阪大学特任教授(兼職)。2019年、大阪電気通信大学情報通信工学部長。博士(工学)。