



# 特集 ライダーデータ利用 I (エアロゾル・雲)

# 衛星搭載ライダーエアロゾル観測のデータ同化

関山 剛\*1

1気象庁気象研究所(〒305-0052 茨城県つくば市長峰 1-1)

Data assimilation of satellite-borne lidar aerosol observations

Tsuyoshi Thomas Sekiyama \*1

<sup>1</sup>Meteorological Research Institute, 1–1 Nagamine Tsukuba, Ibaraki 305–0052

(Received December 1, 2022; revised January 15, 2023; accepted February 08, 2023)

Data assimilation is a mathematical method based on Bayesian statistics for estimating the state of complex systems such as the atmosphere and oceans using observational data and numerical models. While the accuracy of numerical models has been recently improved, the model state can be very different from the real Earth after a few days or weeks of integration. Data assimilation can correct it. Satellite-borne lidars can provide a huge amount of observational data for atmospheric aerosol data assimilation, although the observations are only vertical with a 0-degree view angle. Lidar vertical observations can be extended to horizontally distributed information by data assimilation, in which aerosol concentrations and emissions are optimally estimated. Therefore, the combination of satellite-borne lidar observations and data assimilation is ideal for atmospheric aerosol research. We hope that the atmospheric aerosol products estimated by the combination will be used for global environment monitoring and climate change prediction.

キーワード:データ同化, エアロゾル, ライダー, 衛星観測 **Key Words**: Data Assimilation, Aerosol, Lidar, Satellite Observations

# 1. はじめに

# 1.1 データ同化とは

データ同化とは複雑な系の状態を観測データと数値モデルを使って推定する数理的手法である.その数理 はベイズ統計学に基づく.「同化」という用語は様々な学術分野で使われており,例えば生物学においては 光合成のことを炭酸同化作用と呼ぶ.体内に取り入れた炭酸(=二酸化炭素)を原料にして有機物を合成 し,その有機物を自らの体の一部に組み込むことから「同化」という言葉が使われる.データ同化は観測 データと数値モデルのそれぞれの情報を統合し,現実とよく一致する大気や海洋の状態を計算機で再現す る.なお,データ同化と同じ数理的手法は地球科学以外の分野でも使われている.例えば航空機表面やパイ プライン内の流体の推定,あるいは人工衛星や自動車(カーナビゲーション)など移動物体の位置推定であ る.これらはすべて共通して「誤差を含む限られた量の観測データを元にして,系の真の状態を推定するこ と」を目的としている.

計算機性能の向上に伴い,気象予報モデルや大気化学モデルなど数値モデルの精度は改善を続けている. 数値モデルは地球上の任意の場所や時間における様々な物理量を計算することが可能であり,大変便利である.しかし計算機の中の世界では,数日あるいは数週間も積分を続けると現実の地球とは大きく異なる状態



になってしまう.なぜなら数値モデル自体に物理現象のパラメータ近似や離散化(格子化)にともなう誤差 や数値解法誤差があり,さらに大気の持つ決定論的カオス性が原因で初期値誤差は指数関数的に増大するか らである(ただしそれは瞬間値を再現する場合であり,気候値の再現性は除く).一方で観測は現実の地球 を知る手段であり,誤差の範囲内で世界の真の状態を我々に教えてくれる.しかし任意の場所や時間におけ るあらゆる物理量を観測することは不可能である.現実には地球の巨大さに比べて人間が観測できる範囲は あまりに小さい.これら数値モデルと観測データという相補的な二種類の情報を統合できるのがデータ同化 の利点である.

もともと歴史的にデータ同化は天気予報の数値シミュレーションにおいて予測初期値を手に入れるための 手段として技術開発が進められた<sup>1,2)</sup>.時々刻々推定される大気の状態,すなわち格子点上の気温・気圧・ 風向風速・密度・水蒸気量などは客観解析値と呼ばれ,数値予報の初期値として使われる.昨今の天気予報 の精度を左右するのは数値モデルの性能だけでなくデータ同化の優劣と言っても過言ではない.一方で客観 解析値は気象学の基礎研究において大気全体の状態を広範囲かつ正確に(観測の存在しない物理量も含め て)把握する貴重なデータセットである.エアロゾルのような大気微量成分においてもこのような客観解析 値が手に入れば大変便利である.そこで 2000 年代後半頃から急速に,気象予報技術を活用することで,大 気化学分野においてもデータ同化が広く使われるようになってきた<sup>3,4,5)</sup>.

大気化学分野においては、大気中のエアロゾル濃度を知りたいだけでなく、その地表面放出量や沈着量と いった側面境界情報(=放出・沈着の位置や強度)も重要である。便利なことにデータ同化ではそのような 数値モデルの側面境界情報を統計数理的に推定することも可能である。エアロゾルのような大気汚染物質の 放出量推定は社会的ニーズが高く、重要な研究分野となっている。しかし大気汚染物質の放出量を広い範囲 で直接測定することは不可能であり、その大気中濃度を広範囲に測定することですら困難を伴う。しかし データ同化を使えば、ライダー観測や衛星イメージャー観測のようなリモートセンシングによる濃度値以外 の物理量観測データを使ってエアロゾル放出量の広範囲推定が可能になる。これはいわゆる「逆推計」の一 種である。また、データ同化によって得られた大気エアロゾルの情報を気象予報モデルや気候予測モデルに インプットすることで予報・予測計算の精度を向上させることも可能である。エアロゾルのデータ同化は 様々な分野への貢献が期待できる重要な技術と言える。

#### 1.2 ライダー観測とデータ同化

雲,エアロゾル及び大気成分を人工衛星によって観測する場合を考える.分光計・サウンダー・イメージャーを使って吸収・放射スペクトル線を水平方向に広く走査しながら測定するか、ライダー・レーダーを使って奥行きの情報を含めて一点方向に測定するか、大きく二つの方法があるだろう(ただしここでは掩蔽観測を除く).分光計などを使った方法は鉛直分解能が極めて粗い(あるいは全く鉛直分離できない)が、装置の構造上水平走査がしやすいので地球の広い範囲を観測可能である.その視野幅は数百 km あるいは数千 km に及び、得られる画像はいわば風景写真のような体裁となる.一方、ライダーなどは極めて細かな鉛直分解能を誇るが、装置の構造上水平走査が難しく、衛星に搭載する装置では一方向しか観測できないことが普通である.すなわち視野角は0度であり、鉛直方向の1次元情報を時間方向に連続して繋げて描画するため得られる画像は2次元の平面となり、屏風あるいは切断面のような体裁となる.

通常,3次元以上のデータを人の手で解析することは容易ではない.そのため時間平均や鉛直積算,ある いは気象学では東西平均(同じ緯度の気象条件は類似することが多いため)によってデータの次元を2次元 以下に減らすことによりデータを取り扱いやすくする.その点で鉛直積算量が得られるイメージャー観測は 有利であり,切断面しか得られないライダー観測は不利である.イメージャー観測は観測の時点ですでに次 元の圧縮が行われ(鉛直積算量や特定の高度情報のみを取り出し),理解しやすい図を我々に提供している. 一方でライダー観測はそのような次元圧縮をしないし,水平方向の情報も直接的には我々に与えてくれな い.しかし得られる情報量の総和で考えると,膨大な鉛直分布情報を含むライダー観測はイメージャー観測 に劣るものではない.

大気の流れには上流(風上)と下流(風下)があり,その大気の流れを何らかの方法で知っていれば(例 えば数値モデルで推定),ある一瞬の大気の断面図しか手に入らなかったとしても,原理的には過去(上流) や未来(下流)の姿を再現することができる.データ同化による地球大気の状態推定はまさにそのような作



業と等価である. イメージャー観測のように観測の時点で次元圧縮が行われ, エアロゾル濃度が鉛直積算さ れてしまっては, 鉛直分布情報に関してはほとんど何も観測から得ることができない. 一方, ライダー観測 の鉛直分布情報はデータ同化によって水平方向の分布情報に容易に広げることができる. 情報の圧縮が観測 の際に行われていない分, むしろライダー観測の方が好ましいとも言える. データ同化によって衛星軌道の 左右に広がるエアロゾルプルームを再現できるのであれば, ライダー観測の方がイメージャー観測よりも有 用である可能性もある. ライダー観測とデータ同化は大気エアロゾル研究において理想的な組み合わせと言 えよう.

# 2. データ同化の実際

#### 2.1 データ同化手法

データ同化の数学的操作を極めて簡単に述べるなら、重み付き平均と言ってよい. そこでは単純に数値モ デルの結果と観測データを足して2で割るのではなく、より信用できる方の比重を重くする. 信用の度合い は誤差の相対的な大きさで決まる. モデルの変数と観測の変数が同一で、全てのモデル格子点上で観測が実 施されていれば話はそれだけだが、実際には観測が圧倒的に足りない. 言い換えれば、地球大気の状態を精 縦に表現するのに必要な自由度の巨大さに比して観測データの数は極めて少ないため、圧倒的な劣決定問題 を解く必要に迫られる. そこで各物理量自身の誤差(分散)だけでなく異なる物理量の間の共分散も用いて 重み付き平均を計算する. そうすることで離れた場所の観測データや異なるモデル変数の情報が利用可能に なる. たとえ離れていても相関が高い(=共分散が大きい)と想定される観測データの情報は取り込まれる べきだし(例えば風上風下の位置関係)、物理法則で密に関連しているモデル変数間には高い相関(=大き な共分散)が存在するはずである. 具体的な計算方法は次節で概説する.

データ同化(あるいは一般的にベイズ推定)では数値モデルの予測結果を第一推定値あるいは先験情報と 呼ぶ.もちろん数値モデルは完璧ではないので,第一推定値には誤差が含まれる.問題はこの誤差がどの程 度なのか正確には分からない上に時々刻々変化することである.最新のデータ同化手法では場所・時刻・物 理量の種類に応じて第一推定値にどの程度の誤差が含まれるのか推定することでデータ同化の計算精度を向 上させている.一方で観測データにも誤差は存在し,それは大まかに測器誤差と時間/空間代表性誤差に分 けられる.一般には誤差といえば測器誤差を想像することが多いが,測器の精度を無限に上げていっても観 測誤差を無くすことはできない.なぜなら我々が知りたいのは温度計の先端に触れている(あるいはある瞬 間に熱赤外線を偶然発した)微小空気塊の温度ではなく,小さな試料セルの中に偶然入ってきたエアロゾル の濃度でもなく、少なくとも数百メートル四方,場合によっては数百キロメートル四方の空気塊における平 均温度や平均濃度を知りたいのである.このように極小エリアの観測値と空気塊全体の平均値との差を空間 代表性誤差と呼ぶ.時間方向に関しても同様な齟齬は発生し,時間代表性誤差と呼ばれる.データ同化の計 算に必要な観測誤差は測器誤差と時間/空間代表性誤差の和であり,それらの正確な見積りが重要となる.

観測データが圧倒的に足りない状況では、空間的に少々離れた観測データも精度良く情報を利用でき、数 値モデルで予報している変数以外の物理量の観測データも利用できるデータ同化手法が必要となる. そのよ うな要求を満たすアルゴリズムが四次元変分法(4D-Var)やアンサンブルカルマンフィルタ法(EnKF)で ある. これらのアルゴリズムでは観測演算子と呼ばれる技法を用いることで、数値モデルが再現する変数以 外の観測データも同化できる(次節参照). ライダー観測のデータ同化にはこの観測演算子が必須である. ただし 4D-Var や EnKF は計算機負荷が極めて大きいため、気象予報の分野でもその実用化は計算機性能が 著しく向上した 2000 年代以降のことである.

エアロゾル研究のような大気化学分野でデータ同化が広く行われるようになった要因として、気象予報の 分野においてこれらの高精度データ同化手法(4D-VarやEnKF)が実用化されたことが挙げられる.それ以 前の気象予報データ同化(例えば最適内挿法)は計算機負荷が低い代わりに観測演算子は使えず、また観測 の数と密度が小さい領域ではデータ同化精度が著しく低下する欠点があった.その欠点が、気象予報分野に 比べて観測データが圧倒的に少ない大気化学分野でのデータ同化利用を妨げていた.一方で1990年代頃か ら地球観測衛星が続々と打ち上げられ、大気エアロゾルの広範囲観測データが徐々に得られるようになった ことも背景となり、大気化学分野でのデータ同化利用に繋がった.



#### 2.2 アンサンブルカルマンフィルタの導出

ライダー観測で推定できる要素は波長依存の減衰後方散乱係数や消散係数あるいは偏光解消度だけである が、何故データ同化はエアロゾル濃度を確からしく推定できるのか、以下に簡単な数学的説明を行う.数式 の詳細を追う必要はなく、観測演算子の導入というエッセンスを理解して頂ければ幸いである.4D-Var も EnKF も観測演算子を使うことは共通しており、理想条件下では 4D-Var と EnKF の解は一致する.そこで本 節では、著者が専門とする EnKF を使って観測演算子の便利さを紹介する.この節に興味が湧かなければ次 章 (大気エアロゾルデータ同化の結果を紹介) へ読み進んで下さっても問題ない.

まず,1時刻先の状態を予測する時間発展の式を

$$\boldsymbol{x}_{i+1}^{f} = M\left(\boldsymbol{x}_{i}^{a}\right) + \eta \qquad (1$$

と表す. **x**は N 次元の状態ベクトルであり、そのベクトル成分は数値モデルの各格子点上での気象状態や エアロゾル濃度に対応する.下添字は時刻(*i*)を表し、上添字は予報(*f*)あるいは解析(*a*)を表す.*M* は非線形な数値モデル(いわゆる気象予報モデルや大気化学モデル等)であり、η は数値モデルの誤差を表 す.次に、数値モデルの計算する大気状態やエアロゾル濃度から観測可能な物理量を得るための式を

 $\mathbf{y}_i^o = H_i(\mathbf{x}_i) + \boldsymbol{\epsilon} \qquad (2)$ 

と表す.  $y_i^o$ は p 次元の観測ベクトル(1時刻で入手可能な p 個の観測物理量を並べたもの)であり, Hは非線形な観測演算子(観測モデルとも呼ぶ)である.  $\epsilon$  は観測演算子の誤差を表す.最も単純な観測演算子の例は、数値モデルの格子点上の変数の再現であり、その場合 Hは単に変数を状態ベクトルから抜き出すだけである.もし観測地点が格子点上にないならば、Hは空間内挿も行う.状態ベクトルに含まれない物理量が観測される場合は(ライダー観測はまさにその好例), Hは物理法則に基づく変換演算を行う.要するに、気象状態やエアロゾル濃度の値が鉛直方向にすべて手に入ったとして( $x_i^f$ が該当), どのような減衰後方散乱光や消散が観測されるはずなのか計算するのである.この計算では逆問題を解く必要がないため、計算作業としては比較的シンプルである.この観測演算子の存在がデータ同化において大きな利便性をもたらすことを以下に示す.

カルマンフィルタでは、予報値(すなわち第一推定値) $\mathbf{x}_i^f$ と観測値 $\mathbf{y}_i^a$ から、カルマンゲインと呼ばれる 係数  $\mathbf{K}_i$ を使った重み付き平均で次式のように解析値 $\mathbf{x}_i^a$ を求める.

 $\boldsymbol{x}_{i}^{a} = \boldsymbol{x}_{i}^{f} + \mathbf{K}_{i} \left( \boldsymbol{y}_{i}^{o} - H_{i} \left( \boldsymbol{x}_{i}^{f} \right) \right)$ (3)

ここでカルマンゲイン  $\mathbf{K}_i$  は  $N \times p$  行列であり、 $\mathbf{y}_i^o - H_i(\mathbf{x}_i^f)$  は観測と予報の差を表す p 次元ベクトルである. 加重平均の重みであるカルマンゲイン  $\mathbf{K}_i$  は解析値  $\mathbf{x}_i^a$  の誤差が最小となるように決めるべきである. そこ で、人類が知り得ない地球大気の真値を  $\mathbf{x}_i^t$ とすると、解析値の誤差  $\delta \mathbf{x}_i^a$  は

$$\begin{split} \delta \mathbf{x}_{i}^{a} &= \mathbf{x}_{i}^{a} - \mathbf{x}_{i}^{t} \\ &\approx \mathbf{x}_{i}^{f} - \mathbf{x}_{i}^{t} + \mathbf{K}_{i} \left( \mathbf{y}_{i}^{o} - H_{i} \left( \mathbf{x}_{i}^{t} \right) - \mathbf{H}_{i} \left( \mathbf{x}_{i}^{f} - \mathbf{x}_{i}^{t} \right) \right) \\ &= \delta \mathbf{x}_{i}^{f} + \mathbf{K}_{i} \left( \delta \mathbf{y}_{i}^{o} - \mathbf{H}_{i} \delta \mathbf{x}_{i}^{f} \right) \\ &= (\mathbf{I} - \mathbf{K}_{i} \mathbf{H}_{i}) \, \delta \mathbf{x}_{i}^{f} + \mathbf{K}_{i} \delta \mathbf{y}_{i}^{o} \end{split}$$
(4)

となる. ここで  $\delta \mathbf{x}_{i}^{f} = \mathbf{x}_{i}^{f} - \mathbf{x}_{i}^{t}$  は予報値の誤差,  $\delta \mathbf{y}_{i}^{o} = \mathbf{y}_{i}^{o} - H_{i}(\mathbf{x}_{i}^{t})$  は観測値の誤差である.  $\mathbf{H}_{i}$  は観測演算子 H を $\mathbf{x}_{i}^{f}$ の周りで線形化した行列である. なお, 観測演算子 H が線形ならば(4)式は厳密に成り立つ. しかし非線形であっても通常はかなりの精度で(4)式は成り立つ. この時, 解析値の誤差共分散  $\mathbf{P}_{i}^{o}$  は

$$\begin{aligned} \mathbf{P}_{i}^{a} &= \left\langle \delta \mathbf{x}_{i}^{a} (\delta \mathbf{x}_{i}^{a})^{T} \right\rangle \\ &= \left\langle \left( (\mathbf{I} - \mathbf{K}_{i} \mathbf{H}_{i}) \, \delta \mathbf{x}_{i}^{f} + \mathbf{K}_{i} \delta \mathbf{y}_{i}^{o} \right) \left( (\mathbf{I} - \mathbf{K}_{i} \mathbf{H}_{i}) \, \delta \mathbf{x}_{i}^{f} + \mathbf{K}_{i} \delta \mathbf{y}_{i}^{o} \right)^{T} \right\rangle \\ &= (\mathbf{I} - \mathbf{K}_{i} \mathbf{H}_{i}) \, \mathbf{P}_{i}^{f} (\mathbf{I} - \mathbf{K}_{i} \mathbf{H}_{i})^{T} + \mathbf{K}_{i} \mathbf{R}_{i} \mathbf{K}_{i}^{T} \end{aligned}$$
(5)



と表すことができる.ここで  $\mathbf{P}_{i}^{f} = \left\langle \delta \mathbf{x}_{i}^{f} \left( \delta \mathbf{x}_{i}^{f} \right)^{T} \right\rangle$  は背景誤差共分散,  $\mathbf{R}_{i} = \left\langle \delta \mathbf{y}_{i}^{o} \left( \delta \mathbf{y}_{i}^{o} \right)^{T} \right\rangle$  は観測値の誤差共分散 である.記号 〈 〉は期待値を表す.また,予報値の誤差と観測値の誤差の間には相関がないと仮定した(この仮定は実際にほぼ成り立つ).

誤差がガウス分布に従う場合,最小分散推定によって最尤値を得ることができるから,解析誤差共分散  $\mathbf{P}_{i}^{a}$ の対角和  $tr(\mathbf{P}_{i}^{a})$ の最小値を与える  $\mathbf{K}_{i}$  が求めるべき重みである. そこで対角和  $tr(\mathbf{P}_{i}^{a})$ を  $\mathbf{K}_{i}$  で偏微分する.

$$\frac{\partial tr\left(\mathbf{P}_{i}^{a}\right)}{\partial \mathbf{K}_{i}} = \frac{\partial tr\left(\mathbf{I} - \mathbf{K}_{i}\mathbf{H}_{i}\right)\mathbf{P}_{i}^{f}\left(\mathbf{I} - \mathbf{K}_{i}\mathbf{H}_{i}\right)^{T}}{\partial \mathbf{K}_{i}} + \frac{\partial tr\left(\mathbf{K}_{i}\mathbf{R}_{i}\mathbf{K}_{i}^{T}\right)}{\partial \mathbf{K}_{i}}$$
$$= -2\left(\mathbf{I} - \mathbf{K}_{i}\mathbf{H}_{i}\right)\mathbf{P}_{i}^{f}\mathbf{H}_{i}^{T} + 2\mathbf{K}_{i}\mathbf{R}_{i} \tag{6}$$

ここで $\partial tr(\mathbf{ABA}^T)/\partial \mathbf{A} = \mathbf{A}(\mathbf{B} + \mathbf{A}^T)$ および $\partial tr(\mathbf{AB})/\partial \mathbf{A} = \mathbf{B}^T$ という公式を使い、 $\mathbf{P}_i^f$ および  $\mathbf{R}_i$ が対称行列であることを利用した. (6)式の偏微分を0と置き、 $\mathbf{K}_i$ について解くと

 $\mathbf{K}_{i} = \mathbf{P}_{i}^{f} \mathbf{H}_{i}^{T} \left( \mathbf{H}_{i} \mathbf{P}_{i}^{f} \mathbf{H}_{i}^{T} + \mathbf{R}_{i} \right)^{-1}$ (7)

が得られる.これが求めるべきカルマンゲイン **K**<sub>i</sub>である.(7)式を(5)式に代入すると,解析誤差共分散を 計算するために必要な次式が得られる.

 $\mathbf{P}_i^a = (\mathbf{I} - \mathbf{K}_i \mathbf{H}_i) \, \mathbf{P}_i^f \qquad (8)$ 

これらの式を見ると判るとおり、カルマンゲイン K<sub>i</sub>あるいは解析誤差共分散 P<sup>a</sup><sub>i</sub>を計算するには背景誤差共 分散 P<sup>f</sup><sub>i</sub> =  $\left\langle \delta \mathbf{x}_{i}^{f} \left( \delta \mathbf{x}_{i}^{f} \right)^{T} \right\rangle$ が必要である。そこで、アンサンブル予報シミュレーションを実施し、各アンサンブ ルメンバーのアンサンブル平均からの差(アンサンブル摂動)を統計サンプルと見なして、背景誤差共分散 P<sup>f</sup><sub>i</sub> =  $\left\langle \delta \mathbf{x}_{i}^{f} \left( \delta \mathbf{x}_{i}^{f} \right)^{T} \right\rangle$ の近似値を計算することでこの問題を解決する。アンサンブルメンバーの数を  $m(m \ll N)$ 、 アンサンブルメンバーのサンプル番号を  $k(1 \le k \le m)$ とし、k 番目のメンバーのアンサンブル摂動を  $\delta \mathbf{x}_{i}^{f(k)}$ と すると、誤差共分散 P<sup>f</sup><sub>i</sub>は

$$\mathbf{P}_{i}^{f} \approx \frac{1}{m-1} \sum_{k=1}^{m} \delta \mathbf{x}_{i}^{f(k)} \left( \delta \mathbf{x}_{i}^{f(k)} \right)^{T}$$
$$= \frac{1}{m-1} \delta \mathbf{X}_{i}^{f} \left( \delta \mathbf{X}_{i}^{f} \right)^{T}$$
(9)

と推定できる.ここで  $\delta \mathbf{X}_i^f$ は各列が各メンバーのアンサンブル摂動を成分とする  $N \times m$  行列である.そして (9)式を(7)式に代入すればカルマンゲイン  $\mathbf{K}_i$  が得られ,それを(3)式に代入することで解析値  $\mathbf{x}_i^a$  が手に入る.なお,(9)式を(8)式に代入すると

 $\delta \mathbf{X}_{i}^{a} \left( \delta \mathbf{X}_{i}^{a} \right)^{T} = \left( \mathbf{I} - \mathbf{K}_{i} \mathbf{H}_{i} \right) \delta \mathbf{X}_{i}^{f} \left( \delta \mathbf{X}_{i}^{f} \right)^{T}$ (10)

となる. EnKFによるデータ同化サイクルでは解析値  $x_i^a c$ 得るだけでなく,次のサイクルにおいてアンサン ブル予報シミュレーションを実行するために初期値(すなわち解析値)のアンサンブル摂動  $\delta x_i^{a(k)} (1 \le k \le m)$ が必要である. このアンサンブル摂動を得るための作業をアンサンブル・アップデートと呼び,そのための計算式が(10)である. アンサンブルカルマンフィルタには様々な流儀が存在するが,そのほとんどは(10)式を  $\delta x_i^{a(k)}$ について解くための解法の違いに基づく.

解析値すなわちデータ同化プロダクトを計算するのは(3)式であり、データ同化計算の中で観測値を取り 扱うのはこの部分だけである.そこに観測演算子(プログラミング上は一つのサブルーチン)を介入させる ことで、数値モデルが予報しない物理量の観測データを利用可能にする.また、数値モデルで予報する物理 量と観測する物理量の間を結びつける係数がカルマンゲインだと考えることもできる(そのためカルマンゲ インの計算自体にも(7)式のように観測演算子のサブルーチンが必要).その恩恵によりエアロゾルの場合で あれば物質濃度それ自身だけでなく減衰や偏光のような光学的な情報も同化することができる上、天気予報



の場合ならば大気放射輝度や GPS の電波遅延量でさえ同化できる.

# 3. 衛星搭載ライダーによる大気エアロゾル観測とデータ同化

# 3.1 衛星搭載ライダー CALIOP

大気エアロゾルの観測を主な目的としてこれまで打ち上げられた衛星搭載ライダーはそれほど多くない. 世界最初のものは 1994 年にスペースシャトルに搭載された LITE であり、その後は 2006 年に打ち上げられ た CALIPSO 衛星に搭載された CALIOP<sup>71</sup>と 2015 年に国際宇宙ステーションに短期間搭載された CATS のみ である.これらに加えて 2018 年に打ち上げられた欧州宇宙機関(ESA)の風観測衛星 ADM-Aeolus 搭載の ALADIN では、スピンオフプロダクトとしての扱いでエアロゾル消散係数が公開されている、これらの経 緯は西澤ほか<sup>8)</sup>に詳しく解説されている. CALIPSO 衛星はアメリカ航空宇宙局(NASA)とフランス国立宇 宙研究センター(CNES)によって A-train と呼ばれる地球観測衛星群の中の1台として打ち上げられた. A-train に含まれる人工衛星は隊列を組んで高度約 700 km の太陽同期極軌道を1 周約 98 分で周回している. CALIOP は 532 nm(緑色の可視光)と 1064 nm(近赤外線)の波長を用いて水平分解能 333 m/垂直分解能 30 m で観測を行っており、532 nm においては偏光解消度も測定している。異なる波長における後方散乱強 度はエアロゾル粒子の粒径に強く依存して変化する。また、偏光解消度はエアロゾル粒子の形に強く依存す る. そのため、CALIOP が我々に提供する後方散乱係数あるいは消散係数の観測値はエアロゾルの濃度だけ でなくその組成の情報を多く含んでおり、観測演算子を用いたデータ同化によって、そのエアロゾルの組成 推定も可能となる. このように有益な(しかしそのままでは利用範囲の限られている) ライダー観測データ セットを最大限活用するため、関山ほか<sup>5,9,10</sup>は世界に先駆けて CALIOP エアロゾル観測のデータ同化を試 みた. その詳細を次節以降で紹介する.

#### 3.2 濃度分布の推定

エアロゾル濃度の水平分布を推定するために試みた CALIOP 観測データ同化<sup>5)</sup>では,消散係数ではなく減 衰後方散乱係数の 532 nm および 1064 nm 観測値を利用した. CALIPSO 衛星打ち上げ当初は消散係数が正式 なプロダクトではなかったことが主な理由である.また 532 nm においては偏光解消度の観測値もデータ同 化した.減衰後方散乱係数のデータ同化では,観測演算子によってモデル鉛直層毎の後方散乱強度と往復経 路上での減衰の大きさを推定する必要がある.その推定には数値モデルが予測したエアロゾルの組成・濃 度・粒径分布情報が使われており,その推定値と観測値の差分がカルマンゲインを通じて時空間および物理 変数間で伝達されることによりデータ同化が実現される.偏光解消度については非球形エアロゾル粒子が鉱 物ダスト(いわゆる黄砂)だけであると仮定して,観測演算子によって数値モデルの鉱物ダスト濃度および 粒径分布から偏光解消度の推定値を作成した.また,観測演算子の計算では大気分子による光の散乱や減衰 の影響も考慮してある.

このデータ同化実験では EnKF の実装手法として局所アンサンブル変換カルマンフィルタ(LETKF)<sup>11)</sup>と 呼ばれるアルゴリズムを利用した.また、大気エアロゾル数値モデルとして気象研究所全球エアロゾル予測 モデル(MASINGAR)<sup>12,13)</sup>を使った.MASINGAR は気象庁の現業黄砂解析予測にも使われている数値モデ ルであり<sup>13,14)</sup>,気象研究所の地球システムモデルにもモジュールの一つとして組み込まれて将来気候予測 に役立てられている<sup>15)</sup>.MASINGAR では鉱物ダスト(黄砂)粒子以外にも硫酸塩粒子・黒色炭素粒子・有 機炭素粒子・海塩粒子の挙動を計算しており、観測演算子はそれらすべてのエアロゾル種別の影響を個々に 計算したのち合算している.関山ほか<sup>5)</sup>は計算機負荷を抑えるためMASINGAR の水平解像度を 2.8 度(約 300 km)に設定し、EnKF のアンサンブルメンバー数は 20 個のみにした.この原稿を執筆している 2022 年 現在では計算機性能が向上し、この解像度とアンサンブルメンバー数は見劣りするものとなってしまってい るが、それでもこのデータ同化実験は想定以上に成功した.その結果の一例を Fig.1 で示す.

この例では西日本に大規模な黄砂が飛来した日を図示しており, Fig. 1a と 1b の赤い丸はこの日黄砂を目 視観測した気象官署の位置,青い丸は黄砂が観測されなかった気象官署の位置を示している.黄砂が西日本 を中心に飛来したことが赤色と青色の気象官署の分布で明瞭に判別できる.一方,データ同化を行わずに数 値モデルだけで推定した鉱物ダスト(黄砂)エアロゾルの地上濃度が Fig. 1a であり, CALIOP 観測データ 同化によって推定した地上濃度が Fig. 1b である.前者では本州・四国・九州の広範囲が黄砂に覆われてお





Fig. 1 Simulated surface dust aerosol concentrations on 28 May 2007 (contours and gray shades; daily mean) and the weather stations (red circles) that observed aeolian dust events on the same day. Blue circles indicate the weather stations that did not observe any aeolian dust events on the day. (a) Reference model run without assimilation and (b) CALIOP assimilation results. (c) MODIS optical depth measured on 28 May 2007, in which red shades indicate relatively high values, and yellow shades indicate relatively low values of the optical depth. This figure is taken from the reference<sup>5)</sup>.



Fig. 2 Comparison of lidar observed and simulated extinction coefficients for non-spherical particles (dust aerosol) at 532 nm  $[m^{-1}]$  at the Matsue observatory in western Japan (133°E, 35°N), in May 2007. The X-axis shows the date, in which each tick grid line indicates 00:00UTC. The Y-axis shows the altitude (km). (a) NIES lidar measurements, (b) reference model run without assimilation, and (c) CALIOP assimilation results. Red or yellow shades indicate relatively high values, and blue or blue-gray shades indicate relatively low values of the extinction coefficients. This figure is taken from the reference<sup>5)</sup>.



り、気象官署観測と矛盾している.しかし後者ではデータ同化による黄砂の推定範囲と気象官署観測の分布 がかなりよく一致している.また独立検証データとして、データ同化とは無関係な衛星観測データとして Moderate Resolution Imaging Spectroradiometer (MODIS) によって測定されたエアロゾル光学的厚さ(地上か ら大気上端までの消散係数を鉛直積分した値)を Fig. 1c に示す.このエアロゾル光学的厚さには鉱物ダス ト(黄砂)以外のエアロゾルの影響も含むが、この日は黄砂の影響が支配的であったため光学的厚さの大小 で黄砂プルームの位置が特定可能である.この Fig. 1c でも黄砂は朝鮮半島から九州を経て太平洋に広がっ ており、その分布は Fig. 1a よりも Fig. 1b に(すなわちデータ同化による推定結果に)近い.ライダーで得 られた視野角 0 度かつ減衰後方散乱係数という間接的な情報からのデータ同化によって地上濃度の広範囲な 推定値が得られた好例である.

この時の上空の非球形エアロゾル分布を示したのが Fig. 2 である. 縦軸が高度, 横軸が日時であり, 場所 は国立環境研究所が地上ライダーを設置している島根県松江市である. そのライダーによって観測した消散 係数が Fig. 2a であるが, この観測値はデータ同化には全く使われていない. そして数値モデルだけで推定 した消散係数が Fig. 2b, CALIOP 観測データ同化によって推定した消散係数が Fig. 2c である. 2007 年 5 月 26 日に高度約 2 km のプルームが到達し, その後急速に消散係数(おおよそ濃度に比例する)が減少してい ることが観測値と CALIOP 観測データ同化では示されているが, 数値モデルのみでの推定ではプルーム分



Fig. 3 Dust aerosol emission during a dust storm from May 21 to May 30, 2007. (a) Reference model result without assimilation, (b) optimized model result with CALIOP data assimilation, and (c) difference between the reference and optimized results. Red (blue) shades in the difference map indicate that the optimized dust emission is stronger (weaker) than the reference dust emission. This figure is taken from the reference<sup>9)</sup>.



布が実際よりも低く広くなってしまっている.また、5月12-18日にも弱いプルームが到達しているが、こ れも数値モデルのみでの推定に比べて CALIOP 観測データ同化による推定の方が濃度も高度も現実に即し ている.なお、CALIPSO/CALIOP が島根県松江市の直上を毎日通過することはあり得ないため、このデー タ同化結果(Fig. 2c)は数百 km あるいは数千 km 離れた場所を通過した CALIPSO/CALIOP の観測情報に 基づいて推定されたものだと考えてよい、データ同化の有用性を示す好例である.

### 3.3 地表面放出量の推定

前述したとおり,データ同化を使えばエアロゾルの地表面放出量の推定も可能である.その際,使う観測 データに地表面放出量が含まれていなくてもよい.実際,地表面放出量の広範囲な直接観測は不可能であ る.大気中のエアロゾル濃度(あるいは濃度と高い相関のある物理量)の観測値があれば,その風上の地表 面放出量もデータ同化によって推定できる.その例が Fig.3 である.ここで使ったデータ同化システムは 3.2 節と同様で,CALIOP によって観測された2波長のエアロゾル減衰後方散乱係数と偏光解消度をデータ 同化している<sup>9)</sup>.推定したのは鉱物ダスト(黄砂)エアロゾルの地表面放出量である.推定は全球を対象と したが,ここで図示しているのは黄砂の放出源であるアジア域のみである.

データ同化を行わずに数値モデルだけで推定した地表面放出量(Fig. 3a)では図中央のタクラマカン砂漠 からゴビ砂漠にかけて比較的一様に黄砂が放出されている.これは数値モデル MASINGAR に含まれている 植生分布データや礫被覆率分布データに実況値ではなく気候値や精度の粗い推定値が使われており,多様性 の少ないことが原因であると考えられる.一方, CALIOP 観測データ同化によって推定した地表面放出量 (Fig. 3b)では地域ごとの放出強度の差が大きい.これら二つの差分を示したのが Fig. 3c であり,タクラマ カン砂漠とゴビ砂漠北部で放出が減少するよう CALIOP 観測データ同化は補正しており,ゴビ砂漠南部で は増加するよう補正していることが分かる.最近の研究で他の数値モデルでもゴビ砂漠北部で黄砂発生を過 大評価,南部で過小評価していることが分かっており<sup>16)</sup>,その研究とも矛盾しない補正結果である.しか し残念なことに黄砂に限らずエアロゾルの地上放出量を広範囲に直接観測することは不可能であるため,こ の CALIOP 観測データ同化も直接観測で検証することは難しい.逆に言うと,直接観測の難しい物理量の 広範囲推定にもデータ同化は有用であるということが大きな利点である.さらに,その際にライダー観測の ような視野角0度かつ濃度自体を表すわけではない間接的な物理量が利用できることはデータ同化の利点を 大きく引き立たせる.

#### 3.4 仮想実験

次にデータ同化を使った特殊な仮想実験を紹介する.データ同化に使う観測値を数値モデル(および観測 演算子)で擬似的に作成し、その疑似観測値をデータ同化することで(ただしデータ同化で使う数値モデル は疑似観測値を作った数値モデルとは異なる),疑似観測値を作るために使った数値モデルの状態を外から 推定することができる.これを観測システムシミュレーション実験(OSSE)と呼ぶ.なぜこのような実験 方法が存在するかというと、天気予報の世界では高額な観測施設や観測衛星を導入する前に天気予報の精度 がその追加観測によってどの程度向上するかOSSEによって事前検証したいからである.疑似観測値は数値 モデルで作成しているため、その数値モデルが作る仮想地球に関する情報はすべて手に入る(例えばエアロ ゾルの地上放出量も完全に知ることが可能).したがってOSSEで黄砂放出量を推定すると、その正確な答 合わせが可能となる.関山ほか<sup>10)</sup>は前述のCALIOP 観測データ同化<sup>5,9)</sup>のシステムを利用し、CALIOPの疑 似観測データを別の数値モデルから作成してエアロゾルライダー衛星観測のOSSEを試みた.その際、現実 世界と同じくCALIPSOが1台だけ飛行している仮想世界に加えて、CALIPSOが4台打ち上げられている 仮想世界(それぞれのCALIPSOは同じ軌道で90度ずつ位相をずらして飛行)についても鉱物ダスト(黄 砂)エアロゾル地表面放出量推定を行った.

その結果が Fig. 4 である. ここでは東アジア全体の黄砂エアロゾル放出量の積算値を時系列で示してあ る. 図中, Nature Run (赤線) が仮想地球の放出量 (すなわち正解値), Free-model Run (緑線) がデータ同 化無しでの推定量 (全く異なるモデルを使っているので放出タイミングは似ているが放出量は完全に異なっ ている), 黒線がデータ同化による推定量である. データ同化によって推定量の精度向上が見られるが, CALIPSO が 1 台しか飛行していない場合は正解値とデータ同化推定量の差が乖離してしまうケースが時折





Fig. 4 Time series of the dust emission totaled in the Asian Dust source region (mainly China and Mongolia) in late May 2007. The dust weights of six size bins from 0.200  $\mu$ m to 3.17  $\mu$ m in diameter were accumulated. Red circles indicate the true values produced by the Nature Run. Green triangles indicate free-model results without data assimilation. Black squares indicate the data assimilation results with the CALIOP data assimilation. Only one satellite was used for the data assimilation in (a), and four satellites were used in (b). This figure is taken from the reference<sup>10</sup>.

見られる(Fig. 4a). それに比べて CALIPSO が 4 台飛行しているとその乖離が大幅に縮小することが分かる (Fig. 4b). OSSE は現実世界の CALIOP データ同化の精度検証に使えるだけでなく(Fig. 4a のケース),現 実では実現不可能な理想条件下(例えば CALIPSO が 4 台という Fig. 4b のケース)での仮想実験も可能にす る.

### 4. おわりに

ここで大気エアロゾルデータ同化の視点から考えたライダー搭載衛星ミッションについて少し述べたい. CALIPSO/CALIOP は幸いにも設計寿命を大きく超えて稼働し続けたが,後継衛星がいまだ存在しない (2022 年 12 月現在). 衛星搭載ライダーが我々にもたらしてくれる大気エアロゾル情報は膨大であり,デー タ同化によってさらにその情報を分かりやすく,時空間的に広げることができる. そこから得られる知見は 地上観測や衛星イメージャー観測では得られないものが多く(特に鉛直プロファイル情報),大気エアロゾ ル観測を主目的としたライダー搭載衛星が上空を飛行し続けてくれることを願ってやまない. その願いを叶 えてくれる衛星ミッションが ESA と我が国の宇宙航空研究開発機構 (JAXA)が共同で進めている Earth Clouds, Aerosols and Radiation Explorer (EarthCARE) である<sup>17,18)</sup>. EarthCARE は度々打ち上げが延期されて きたが,ついに 2023 年に打ち上げられることになった. この衛星に搭載される大気ライダー ATLID につい ては西澤ほか<sup>8,19)</sup>に詳しい. EarthCARE はドップラー雲レーダーも搭載しているため,エアロゾル・雲・対 流・降水プロセスの複合的な観測にも期待が持てる. 我々大気エアロゾル研究者は EarthCARE に大きな期 待を寄せている.

さらに、衛星搭載ライダーの打ち上げ数が限られている中では、エアロゾル観測が主目的ではない衛星搭載ライダーの観測データについても活用が必要だろう. ADM-Aeolus/ALADIN はスピンオフプロダクト扱い



ではあるが消散係数などを公開しており、エアロゾルデータ同化での活用に期待が持てる.また、日本でも 衛星搭載コヒーレントドップラー風ライダー(CDWL)の実現に向けて活発な動きがあり<sup>20,21)</sup>、その実現 を期待したい.天気予報の分野では高層大気の風観測データが圧倒的に不足しているため、その観測値が得 られた場合、天気予報精度向上に与えるインパクトが大きい.高層大気の風速観測は衛星搭載ライダーの得 意とするところであり、天気予報は社会的ニーズが大きいため、大気エアロゾル観測を主目的とする場合よ りも国民の理解を得やすいのではないだろうか.また、JAXA は超低高度衛星(例えば Super Low Altitude Test Satellite; SLATS)の運用技術に大きく長けており、SLATS 後継機へのドップラー風ライダー搭載にも期 待が持てる.そのような衛星が多数打ち上げられるようになれば、3.4節で紹介した仮想実験のように、よ り高精度な大気エアロゾルデータ同化が実現できるだろう.そのように作成された大気エアロゾルプロダク トが地球環境監視や気候変動予測に役立てられる日が来ることを願う.

# 謝 辞

EnKFのソースコードは三好建正博士(元気象庁数値予報課,現理化学研究所),全球エアロゾル数値モ デルのソースコードは田中泰宙博士(元気象庁気象研究所,現気象庁数値予報課)から提供頂きました.ラ イダー用の観測演算子開発には清水厚博士(国立環境研究所)から多大な協力を頂きました.

#### 引用文献

- 1) J. M. Lewis, "Sasaki's Pathway to Deterministic Data Assimilation," in Data Assimilation for Atmospheric, Oceanic and Hydrologic Applications, S. K. Park and L. Xu, ed. Springer, pp. 1–19 (2009).
- I. M. Navon, "Data Assimilation for Numerical Weather Prediction: A Review," in Data Assimilation for Atmospheric, Oceanic and Hydrologic Applications, S. K. Park and L. Xu, ed. Springer, pp. 21–65 (2009).
- K. Yumimoto, I. Uno, N. Sugimoto, A. Shimizu, Z. Liu, and D. M. Winker, "Adjoint inversion modeling of Asian dust emission using lidar observations," Atmos. Chem. Phys., 8, 2869–2884 (2008).
- 4) A. Benedetti, J.-J. Morcrette, O. Boucher, A. Dethof, R. J. Engelen, M. Fisher, H. Flentje, N. Huneeus, L. Jones, J. W. Kaiser, S. Kinne, A. Mangold, M. Razinger, A. J. Simmons, and M. Suttie, "Aerosol analysis and forecast in the European Centre for Medium-Range Weather Forecasts Integrated Forecast System: 2. Data assimilation," J. Geophys. Res., 114, D13205 (2009).
- 5) T. T. Sekiyama, T. Y. Tanaka, A. Shimizu, and T. Miyoshi, "Data assimilation of CALIPSO aerosol observations," Atmos. Chem. Phys., **10**, 39–49 (2010).
- 6) IPCC (ed.), Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press (2021).
- D. M. Winker, H. H. Hunt, and M. J. McGill, "Initial performance assessment of CALIOP", Geophys. Res. Lett., 34, L19803 (2007).
- 8) 西澤智明,神慶孝,石井昌憲,岡本創,"衛星搭載高スペクトル分解ライダーによるエアロゾル・雲の全球観 測,"レーザセンシング学会誌,1(2),55-60 (2020).
- 9) T. T. Sekiyama, T. Y. Tanaka, T. Maki, and M. Mikami, "The Effects of Snow Cover and Soil Moisture on Asian Dust: II. Emission Estimation by Lidar Data Assimilation," SOLA, 7A, 40–43 (2011).
- T. T. Sekiyama, T. Y. Tanaka, and T. Miyoshi, "A simulation study of the ensemble-based data assimilation of satellite-borne lidar aerosol observations," Geosci. Model Dev. Discuss., 5, 1877–1947 (2012).
- T. Miyoshi and S. Yamane, "Local Ensemble Transform Kalman Filtering with an AGCM at a T159/L48 Resolution," Mon. Wea. Rev., 135, 3841–3861 (2007).
- 12) T. Y. Tanaka, K. Orito, T. T. Sekiyama, K. Shibata, M. Chiba, and H. Tanaka, "MASINGAR, a global tropospheric aerosol chemical transport model coupled with MRI/JMA98 GCM: Model description," Papers in Meteorology and Geophysics, 53, 119–138 (2003).
- 13) 田中泰宙,小木昭典,"気象庁全球黄砂予測モデルの更新について,"測候時報,84,109-128 (2017).
- 14) 小木昭典,松本隆則,北島俊行,田中泰宙,弓本桂也,関山剛,眞木貴史,"気象庁全球黄砂解析予測モデルに おけるエーロゾルデータ同化手法の導入について,"測候時報,89(1),1-15(2022).
- 15) S. Yukimoto, H. Kawai, T. Koshiro, N. Oshima, K. Yoshida, S. Urakawa, H. Tsujino, M. Deushi, T. Tanaka, M. Hosaka, S. Yabu, H. Yoshimura, E. Shindo, R. Mizuta, A. Obata, Y. Adachi, and M. Ishii, "The Meteorological Research Institute Earth System Model version 2.0, MRI-ESM2.0: Description and basic evaluation of the physical component," J. Meteor. Soc. Japan, 97, 931–965 (2019).
- 16) T. T. Sekiyama, Y. Kurosaki, M. Kajino, M. Ishizuka, B. Buyantogtokh, J. Wu, and T. Maki, "Improvement in dust storm simulation by considering stone coverage effects for stony deserts in East Asia," J. Geophys. Res., 128, e2022JD037295 (2023).



- 17) 菊池麻紀, "雲エアロゾル放射ミッション (EarthCARE)", 天気, 64, 59-61 (2017).
- 18) 菊池麻紀, 沖理子, 久保田拓志, 吉田真由美, 萩原雄一朗, 高橋千賀子, 大野裕一, 西澤智明, 中島孝, 鈴木健 太郎, 佐藤正樹, 岡本創, 富田英一, "雲エアロゾル放射ミッション「EarthCARE」,"日本リモートセンシング学 会誌, **39** (3), 181-196 (2019).
- 19) 西澤智明,工藤玲,日暮明子,及川栄治,岡本創, "EarthCARE 衛星搭載ライダーデータを用いたエアロゾル・ 雲推定アルゴリズム,"日本リモートセンシング学会誌, **39**(3), 215-224 (2019).
- 20) 石井昌憲, "大気風を診るドップラー風ライダー,"大気化学研究, 40, 040A01 (2019).
- 21)石井昌憲,岡本幸三,久保田拓志,藤平耕一,松本紋子,今井正,境澤大亮,今村俊介,石橋俊之,田中泰宙, 佐藤篤,西澤智明,村田健史,岡本創,沖理子,佐藤正樹,岩崎俊樹,"衛星搭載ドップラー風ライダーによる 全球風観測,"レーザセンシング学会誌,1(2),61-66 (2020).



# 関山 剛

気象庁気象研究所 全球大気海洋研究部 主任研究官. 1995年,大阪大学理学部化学科卒. 同大学 院理学研究科博士前期課程有機化学専攻中退. 1996年,気象庁に入庁. 4年間の気象庁本庁勤務 ののち,2000年に気象研究所へ異動.2005-2006年,米国大気研究センター (NCAR)に出向. 2007年より主任研究官.2013年,東北大学大学院理学研究科地球物理学専攻にて博士(理学) の学位を取得.2015-2016年,鳥取大学乾燥地研究センター客員教授.気象予報士,第1種放射 線取扱主任者,甲種危険物取扱者の資格を所持.専門分野は大気微量成分の数値シミュレーショ ンとデータ同化.ミドルネームはThomas.