

jh250035

人機一体気象モデリングで拓く 次世代メソアンサンブル気象予測システム

澤田洋平（東京大学）

概要

次世代メソアンサンブル気象予測システムの構築により、豪雨などの災害予測とそれに基づいた社会的意思決定の改善を目指す。本年度は多様な事例におけるメソ気象シミュレーションのモデル選択とモデルパラメータ選択手法を確立し、多数事例のデータセット作成と最適化結果を得ることができた。これはモデル選択・モデルパラメータ選択を静的なもののみで最適化する手法だが、これを動的なものにとらえる手法も検討し、次年度以降に時々刻々と得られる観測データを用いて逐次的にオンライン学習する準備ができた。Urgent computing による災害時のみ大アンサンブルを生成する手法については基礎的・理論的な検討を重ねている。これに関連する研究として、大アンサンブルがリアルタイムに利用可能になったことを想定して、リアルタイムに災害の最悪シナリオを解析する手法を提示することができた。

1. 共同研究に関する情報

(1) 共同利用・共同研究を実施している拠点名

東京大学 情報基盤センター

東北大学 サイバーサイエンスセンター

①気象シミュレーション不確実性定量化(研究項目 A,B)、②極端気象予測の urgent computing (研究項目 C)、③統合ソフトウェア開発

(2) 課題分野

大規模計算科学課題分野

2. 研究の目的と意義

気候変動の進行により大規模な気象災害は増加傾向にある。気象災害の発生を適切に予測し、国・自治体・市民の全てのレイヤーで適切な防災・減災行動を促す技術が希求されている。近年この分野は研究者のプロセス理解に基づき構築された物理ベースのモデルの出力の一部ないし全部を学習して AI で置き換えるような「人機一体気象モデリング」と呼ぶべき戦略が主流となり、全球天気予報の精度は飛躍的に向上している。

しかし、洪水をもたらすような豪雨予測は未だ難しい。地域レベルの豪雨を確率的に予測する基幹技術として、地域レベルの大気現

(3) 参加研究者一覧と役割分担

- ・澤田洋平（東大工）（代表）①
- ・中島研吾（東大情基セ）（副代表）②③
- ・奥川伸一、日比野研志、南出将志、Le Duc、Islam Md. Rezuanul、（東大工）、Namal Rathnayake (JAMSTEC)、橋本和宗（阪大工）①
- ・本田匠（東大情基セ）②
- ・住元真司・埜敏博・荒川隆・下川辺隆史（東大情基セ）、芝隼人（兵庫県立大学）・河合直聡（東北大学）、八代尚（国環研）、松葉浩也（日立製作所）③

象を高精細に解くメソ気象モデルを多数回実行するメソアンサンブル気象予測システム(Mesoscale Ensemble Prediction System: MEPS)がある。MEPS は日本の気象庁をはじめ、ベトナムやバングラデシュといった開発途上国でも幅広く運用され、極端気象の適切な予測や、多数のシミュレーションを用いたシナリオ分析などへ利用が進む。現状のMEPSが抱える問題点は、以下の3点である。

問題点①: 地域のレアイベントをターゲットとするという MEPS の性質上、そのすべてを AI で置き換えるだけの十分な学習データを得ることが難しい

問題点②: 適切なアンサンブルを生成するためには気象モデルに内在する様々な不確実性を定量化する必要があるが、例えば気象庁の MEPS は初期値の不確実性のみに着目し、気象モデルのモデル選択やパラメータ選択にまつわる不確実性を考慮していない。

問題点③: 高精細なメソ気象シミュレーションを多数回実行する MEPS は計算コストが高いが、極端イベントを適切にとらえるには大きなアンサンブル数が必要で、そのリアルタイム実行は難しい。

本研究の目的は、「人機一体気象モデリング」の技術トレンドに立脚し、上述した①～③の状況に適した次世代 MEPS のアルゴリズム提案・実装・ソフトウェア基盤構築を行うことである。具体的には以下の3つの機能を開発し、それらを有機的に繋ぐことでMEPSに革新をもたらすことを目指す。

機能 A: 気象モデルに内在する不確実性を定量化するための過去の観測データを用いたオフライン学習機能

機能 B: オフライン学習の結果を事前分布としつつ、リアルタイムに得られる新しい観測データを取り込み状態推定を行いながらその場その場でモデルパラメータなどを調整するオンライン学習機能

機能 C: 極端気象の発生が予見されるとき

み生成 AI で大アンサンブルの初期値を生成し大規模計算を行う Urgent Computing の機能

いずれの機能開発においても極端気象の物理的理解に立脚したシミュレーションに関する気象学、プロセス駆動シミュレーションとデータ駆動モデリングを組み合わせることで不確実性定量化・データ同化・人工データ生成を行う統計数理、そして最新鋭の計算機において「シミュレーション+データ+学習」の融合を目指す計算科学が協働する必要がある。開発を通じて、極端気象の予測という特定の問題にとどまらない数理的・計算科学的な知見を得ることも本研究の重要な目的である。

3. 当拠点の公募型共同研究として実施した意義

●本研究は気象学・数値天気予報技術に立脚した研究課題でありつつも、単にプロセスを直接記述するようなモデリングのあり方を変革し、観測データに基づいたモデリングとの融合を目指すものである。それを目指すために「シミュレーション+データ+学習」という計算科学的な挑戦を含んでいる。またモデルに内在する不確実性を推定するといった場面では統計数理や AI の知見が重要である。このような研究課題の学際性を鑑みて、共同研究として実施することに意義があった。

●本研究では、物理プロセスの理解に基づいて作られた気象シミュレーションの超並列実行と、その実行結果の巨大データの機械学習による解析を連携して行う必要がある。従って東京大学情報基盤センターの Wisteria/BDEC01 が有するシミュレーションノード群とデータ・学習ノード群の双方の資源利用が必要であった。

4. 前年度までに得られた研究成果の概要

該当なし。

5. 今年度の研究成果の詳細

研究項目 A: 気象シミュレーションの不確実性 定量化とオフライン最適化

前年度まで実施した JHPCN 課題「極端気象現象予測の不確実性の起源の解明」で 1000 メンバを超える大アンサンブル計算の実行とこの結果を機械学習で置き換えるサロゲートモデルにより、モデル選択とパラメータ選択の不確実性定量化・最適化を実現している。現状ではこの最適化は豪雨に特化している。豪雨以外の高温・低温などの極端イベントにも対応すると同時に、豪雨時以外にでも安定してモデルが動くよう汎用性を高める開発を行った。

まず、豪雨以外の熱中症を引き起こす高温などの極端イベントにおいては気象モデルにおいて陸面を扱うスキームの働きが大きいが、現行のシステムでは陸面に関するスキームの不確実性を定量化できるようにできていなかった。そこで陸面スキームのモデル選択やパラメータをモンテカルロシミュレーションすると同時に、これらを観測データから最適化できるようにシステムの改良を行った。

改良した新システムにおいて、豪雨にとどまらない多様な極端イベントを対象に、1000 アンサンブルの計算をこれまでと同様にベトナム水文気象庁との協力の上で行った。これまでに作成した 1000 アンサンブル計算のデータセットとしては豪雨・高温・低温イベントを網羅することになる。これをこれまでと同様にガウス過程回帰で学習することで、多様なイベントを高性能に予測できるモデル選択を見出し、アンサンブルを構成した。

当初予期していた通り、豪雨をうまく予測できるモデル選択と、低温・高温イベントをうまく予測できるモデル選択は必ずしも一致しない。また低温・高温イベントについてはモデル選択間の等結果性(=equifinality)が大きく、様々なモデル選択の組み合わせが同等程度の性能を持つことが分かった(図 1)。これらの結果からは、モデル選択を単に最適化するだけでは吸収不可能

な系統誤差が気象モデルに存在することを示唆する。現業的には、研究項目 B で開発しているような、時々刻々の状況変化に応じてモデル選択を柔軟に変更するような手法が有効であることがわかる。

豪雨に関する結果は査読付き学術論文として本年度出版済みである(<https://doi.org/10.1175/WAF-D-24-0244.1>)。低温・高温イベントに関する結果も現在論文として取りまとめている。

研究項目 B: 気象シミュレーションのリアルタイム状態推定とオンラインモデルパラメータ学習

前年度まで実施した JHPCN 課題「極端気象現象予測の不確実性の起源の解明」で開発した HOOPE-EnKF を研究項目 A で開発したシステムに適用する。これにより、事前のオフライン学習に加えて、時々刻々と入手できる観測データに応じてその場でモデルパラメータを最適化する。

現在は研究項目 A で用いている改良版の Weather Research and Forecast (WRF)モデルに HOOPE-EnKF を実装する作業を進め、動作確認をしている段階である。

また、全球モデルの解析値(=観測データをモデルにデータ同化した状態量の推定値)を WRF で単にダウンスケーリングする形で予報を行うことを想定する場合(=共同研究を行っているベトナム水文気象庁の現業天気予報の設定)、逐次推定すべきなのはモデル選択とモデルパラメータ選択をつかさどる 5-25 個程度のパラメータであるため、本質的な次元は小さい。この着想から、HOOPE-EnKF の姉妹アルゴリズムで、(高次元問題には対応できないが)より計算複雑性が小さい粒子フィルタベースの HOOPE-PF(<https://doi.org/10.1029/2021MS002882>)が利用できると考え、そちらの実装も進めた。小規模実験を通じた検討では、研究項目 A で得られるパラメータのオフライン推定の確率分布が正規分布ではないために、正規分布を仮定する

HOOPE-EnKF よりも HOOPE-PF の方が優れていることが伺えた。ただし HOOPE-PF では研究項目名に示した「リアルタイム状態推定」は、状態空間が高次元であるために行えないため、引き続き HOOPE-EnKF の開発も継続する。

本年度においては主に研究項目 A で計算量をすべて消費したこともあり、実際の気象データを使った本計算に進むことはできなかった。しかし来年度ベトナム水文気象庁現業と同じ問題設定で HOOPE-PF によりモデル選択の組み合わせを動的に推定する枠組みを適用できる準備が整った。

最後に、気象モデル WRF は米国で開発された気象モデルだが、近年あまり積極的な開発が行われなくなっている。将来的な拡張可能性なども鑑みて、WRF での開発と並行して、日本で広く開発が進められている気象モデルである SCALE での開発も並行して進めている。SCALE は WRF と比べても可読性に優れたアンサンブルカルマンフィルタのコードが実装されていて(SCALE-LETKF)、これを有効活用してモデルパラメータのオンライン・オフライン学習機能の実装を進めた。こちらはまだ実際の気象データに適用できるだけの準備は整っていないため、来年度も開発を続ける。

研究項目 C: 極端気象の Urgent computing 大アンサンブル予測

モデル誤差や真値を制御した理想化シミュレーション実験に向け、簡易な全球大気モデルに基づくデータ同化システム SPEEDY-LETKF の動作を確認した。機械学習フレームワーク anemoi を採用し、SPEEDY-LETKF の出力を用いた学習と推論を行うために必要な開発を行なった。十分な推論精度が得られていないため、モデル誤差などの条件について見直しを進めている。

関連する研究として、現業での天気予報を想定して、大アンサンブル予測をリアルタイムに処理して、各地の災害ハザードを分析する手法の

開発を行った。台風を事例として、台風が引き起こす 3 つの災害(高潮・洪水・風害)のハザードを 1000 個のアンサンブルすべてについて計算した。高潮と風害については既往のパラメトリックモデルを用いた。洪水に関しては日本の気象庁と協力し、流域雨量指数(通称:キキクル)の計算に用いられている現業モデルを用いた。

リアルタイムに災害のリスクを算定するためには計算した 1000 個すべてのアンサンブルを一覧するだけでは不十分である。そこで実務上重要な「最悪シナリオ」の算出に取り組んだ。複数地点における複数の災害を考慮したとき、「最悪シナリオ」を 1 つに絞り込むことは困難である。そこで、多目的最適化で用いられるパレート最適性の考え方を用い、パレートフロンティアとして最悪なアンサンブルの群を抽出した。そのうえで、この最悪シナリオ群を Affinity propagation によってクラスタ分類し、最終的に 5-10 個程度の最悪シナリオ群を特定した。パレートフロンティアの計算もクラスタ分類も計算コストは極小であり、リアルタイムに 1000 アンサンブルから高々数個の最悪ハザードを有する代表的なシナリオを算出することができる。

このシナリオ分析の成果は本年度論文として出版されている(<https://doi.org/10.1175/WAF-D-25-0030.1>)。本研究項目が主眼としている 1000 メンバーアンサンブルの高速生成が実現すれば、その下流タスクはすでに解かれており、「1000 メンバは生成できるけれどその利用法がよくわからない」といった事態にはならない。

6. 進捗状況の自己評価と今後の展望

本年度は 3 年計画の 1 年目である。研究項目 A については、ほぼ完了したとあってよい。FY2026(2 年目)まで多様な観測、多数事例での最適化を続けるという計画になっているが、すでにそのような最適化を行い、5 章で記したような知見を得ることができた。当初計画よりも順調に進展し、FY2026(2 年目)は研究項目 B, C に集

中することができる。

研究項目 B については、研究項目 A の成果を拡張し、時々刻々と入ってくる観測データからオンライン学習でモデル選択・モデルパラメータ選択を最適化するということを目指し、開発を行った。当初計画ではモデルの状態量も同時推定できる HOOPE-EnKF の実行を目指したが、モデルに内在する誤差のみを扱う (i.e., モデル選択・モデルパラメータ選択のみを解く) 戦略も十分有効でかつ実装しやすいと考え、そのような方向性での検討も進めた。FY2026(2 年目)は実際の性能検証を予定しており、そこに向けた準備が整っているという点では順調に進展している。

研究項目 C はトイモデルや低解像度気象モデルでの検討を続けている。既往研究でも示されている通り、物理モデルによるアンサンブル計算を学習して、大量のアンサンブル初期値を生成すること自体は可能である。しかし、そのアン

(次ページに図表あり)

サンブルが本来あるべき物理的一貫性や適切な誤差構造情報を保持した、「良いアンサンブル」であることが重要で、そのような良いアンサンブルを少ないアンサンブル数でのシミュレーション結果のみから作るためにさらなる検討が必要である。計画通り FY2026(2 年目)において実際のメソ気象モデルの出力を用いた学習と生成に移れるかは少し不安材料となるが、問題の所在は特定できており、研究活動が持つ不確実性の範囲内で研究を進行できていると考えている。

研究項目 C については、大アンサンブル計算が実現した後の高速なシナリオ分析の困難さについてはある程度技術的に克服することができている。この点は順調に研究が進展している。

※7.研究業績はウェブ入力です

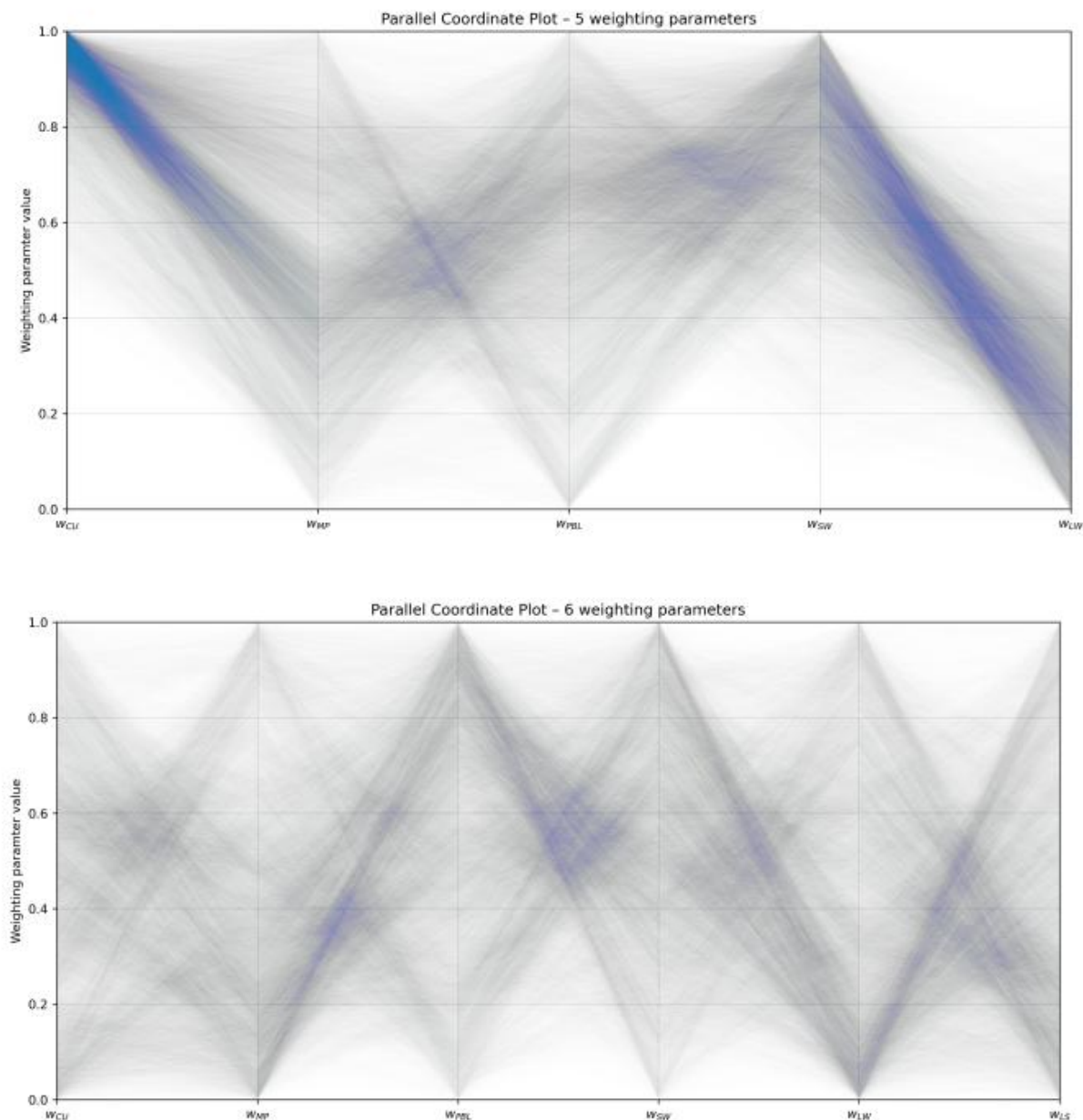


図 1. 最適化後の事後分布におけるモデル選択の重みパラメータの図示。横軸がスキームを示しており、左から積雲対流、雲美物理、境界層、短波放射、長波放射、(下図のみ) 陸面過程を示す。0 と 1 の端が二つの候補モデルのうち片方のみを選択することを示し、その間は二つの候補モデルを毎時間タイムステップ重み付け平均して時間発展させることを示す。一本の線が一つの重みパラメータの組を示している。(上) 豪雨事例で最適化した結果、(下) 低温事例で最適化した結果。事例が異なると最適なモデル選択が大きく異なることがわかる。また豪雨事例の方がパラメータの決定性が大きく、低温事例では様々な異なる組み合わせのモデルが同じように誤差を低減できることが分かった。