

jh220020

極端気象現象予測における不確実性の起源の解明

澤田洋平（東京大学）

概要

非線形性の強い極端気象の予測における不確実性がどのようなプロセスで生じるかを明らかにすることで極端気象予測の性能改善を目指す。大規模気象シミュレーションの初期値およびモデルパラメータ、さらにはモデル選択に起因する予測の不確実性を効率よく定量化する統計数理的手法を開発した。また、その手法を Wisteria/BDEC-01 上で効率よく動作させるための統合ソフトウェア開発も行った。開発したソフトウェアの適用により、日本域に接近する大量の台風の予測に対する不確実性の起源を網羅的に解析するための計算基盤の構築も進めた。以上の成果により来年度以降 Wisteria/BDEC-01 の性能を引き出したうえで、上記の目的に迫る計算を行う準備が整った。

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（該当するものを残す）

東京大学 情報基盤センター

- (2) 課題分野（該当するものを残す）
大規模計算科学課題分野

- (3) 共同研究分野（HPCI 資源を利用している研究課題のみ、該当するものを残す）
超大規模数値計算系応用分野
超大規模データ処理系応用分野
超大容量ネットワーク技術分野
超大規模情報システム関連研究分野

- (4) 参加研究者の役割分担

- ・ 澤田洋平（東大工）（代表）①
- ・ 中島研吾（東大情基セ）（副代表）④
- ・ 南出将志（東大工）②

・ Le Duc、Islam Md Rezuani（東大工）、橋本和宗（阪大工）③

・ 住元真司・埴敏博・河合直聡・荒川隆（東大情基セ）、八代尚（国環研）、松葉浩也（日立製作所）④

注：①不確実性定量化、②気象データ作成、③気象データ機械学習、④統合ソフトウェア

2. 研究の目的と意義

洪水・浸水被害をもたらす極端気象現象の予測には未だ大きな不確実性がある。大規模気象シミュレーションのどこに予測の不確実性をもたらす要因が存在しているのかを網羅的に理解することが重要である。既存の研究では初期値・境界値の不確実性を解析するのにとどまっている。そこで本研究では大規模気象シミュレーションに内在するすべての不確実性を効率よくベイズ推定するソフトウェア基盤の構築を目的とする。また開発したソフトウェアを用いて不確実性情報が付加された大規模な気象データの作成・公開を行い、その

データ解析から極端気象現象予測の不確実性の起源を解明する。

ゲリラ豪雨・線状降水帯・台風といった災害をもたらす気象現象の予測には非線形な雲・対流プロセスが重要である。本研究は最先端気象シミュレーションにおいてこれらのプロセスの予測の不確実性の起源を解明することを目指すもので、気象学における Grand Challenge に挑む意義ある研究である。

シミュレーションの不確実性定量化は、予測精度の向上、物理プロセスの理解、意思決定への貢献などで重要である。既存研究は多いが、気象計算のような大規模シミュレーションに対しては決定版と呼ぶべき手法がない。「シミュレーション+データ+学習」融合を推し進めこれを実現し、気象計算に限らず JHPCN 各センターで行われている大規模シミュレーションに基づく予測研究全体に貢献できる。

3. 当拠点の公募型研究として実施した意義

●本研究課題は、極端気象現象の理解とそのシミュレーションに関する気象学、不確実性定量化に関する統計数理、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」の融合をめざす情報科学が協働する必要がある。この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

なし

5. 今年度の研究成果の詳細

①気象シミュレーションの不確実性定量化手法の開発

時間変化するパラメータも含んでその最適化と不確実性推定を行うことができる新しいデータ同化手法[1]の開発と高度化を進めた。特に[1]の手法では気象シミュレーションのような大自由度系に拡張することは難しかったが、気象学で用いられていることの多いアンサンブルカルマンフィルタからの簡単な拡張で、モデルパラメータの動的な不確実性を定量化する手法を開発した。データ同化手法開発における代表的なテストベッドである Two-scale Lorenz96 モデルを用いた数値実験でその有効性を確認することができた。本成果はすでにプレプリントとしてまとめられている (<https://arxiv.org/abs/2305.07798>; 投稿は 2023 年 5 月)

②気象シミュレーションの不確実性定量化手法の応用

①で開発したソフトウェアの応用先として、台風予測の不確実性解析を行う。日本に付近に接近するすべての台風の微細な 3 次元構造を再現しその予測の不確実性をアンサンブルとして付加した巨大気象データセットを新たに構築し、その機械学習による解析を行う。気象シミュレーションソフトウェアとしては Weather Research and Forecasting (WRF) を使い、静止気象衛星ひまわり 8 号の高時空間分解能観測を WRF にデータ同化することでデータセットを作成する。これまで別の計算機で運用されていたすべてのコード資源の Wisteria/BDEC-01 への移植作業がおおむね完了した段階である。計画書における研究項目 B で説明したデータセットの version 1 を整備する準備はおおむね整っており、今後計算機資源を大規模に利用して、データ生成を進める予定である。

上記の version 1 というのは WRF のモデル化/パラメタリゼーションそのものに内在する不確実性を陽に考慮しない。これを考慮する version 2 の整備に向けて、事例解析を行った。具体的には 2022

年に非常に強いスーパー台風として話題になった台風 Nanmaadol を対象に、WRF の 22 個のパラメータを最適化し、予測性能向上を企図する数値実験を行った。海上で急発達する台風を観測するすべは必ずしも多くない。ここでは静止気象衛星ひまわり 8 号の観測を用いてこの観測から 22 個のパラメータをどこまで拘束できるかを調べた。画像認識の分野で画像間の類似性指標に用いられる Structural Similarity を評価関数に組み入れる独自の手法で、WRF の雲微物理および境界層のパラメータを適切に最適化することに成功した。これにより Nanmadol の強度予測が向上することが確認できた。一方で、ひまわり 8 号の観測のみで不確実性を大きく減らすことのできるパラメータの数は多くなく、より多くの観測を得たり、より多くの台風事例を用いたりしなければ十分に予測の不確実性を減らすことはできないことが分かった。

加えて、作成したデータセットの新しい解析手法に関しても開発を進めている。特筆すべき成果としては経済学のパレート最適性の観点を取り入れたアンサンブルを用いた新しい災害リスク解析手法を開発し、現在論文投稿済み (<https://www.authorea.com/doi/full/10.22541/essoar.167979634.48767251/v1>) である。今後は機械学習等も用いて過去の台風がもたらす災害リスク解析とその予測の不確実性が生まれる起源について理解を深めていく予定である。

③統合ソフトウェア開発

h3-Open-SYS/WaitIO は複数の並列プログラム間で通信を行うためのライブラリである。各並列プログラムは典型的には MPI プログラムであるが、MPI 以外の形の並列プログラムからも利用できるよう汎用的に設計されている。API は MPI 類似でありアプリケーション変換も容易である。

2022 年度は、Socket を利用した h3-Open-SYS/WaitIO-Socket (以下、WaitIO-Socket) 上でのアプリケーションとして h3-Open-UTIL/MP (以下、UTIL/MP) を利用した機械学習 (ML) との融合、Wisteria/BDEC-01 上での異種システム間ジョブ実

行システムの開発、ならびに、幅広い異種システム間結合をサポートするため、共用ファイルシステム上での通信を実現する h3-Open-SYS/WaitIO-File (以下、WaitIO-File)、Socket と File を組み合わせた h3-Open-SYS/WaitIO-Hybrid (以下、WaitIO-Hybrid) を開発した。

WaitIO-Socket と h3-Open-UTIL/MP を用いて Wisteria/BDEC-01 上で気象シミュレーションと機械学習を融合したアプリケーションでの評価において Odyssey 上のみで気象シミュレーションと機械学習を実行した結果よりも Odyssey 上で気象シミュレーション、GP-GPU を搭載した Aquarius 上で機械学習を実行する異種システム連携の方が 35% 高速であることが分かった。

システムワイドな異種連成計算実行環境を国内外の計算機センターでより広く利用して頂くために、TCP/IP だけでなく共有ファイルシステムを用いた (h3-Open-SYS/) WaitIO-File、更に WaitIO-Socket と WaitIO-File を組み合わせた WaitIO-Hybrid を開発した。実現した WaitIO-File、WaitIO-Hybrid の通信性能は Odyssey の WaitIO-Socket の通信性能より高速であることが分かった。また、WaitIO-Socket と WaitIO-File を組み合わせた WaitIO-Hybrid の結果は通信遅延が大きな WaitIO-File の情報交換遅延を WaitIO-Socket により削減し通信性能が向上した。

汎用カプラ h3-Open-UTIL/MP の開発において 2022 年度は h3-Open-SYS/WaitIO と協調した異種間連成の性能評価を実施するとともに、ジョブ実行環境の整備を行った。

WaitIO を用いた異種間連成では機種毎にジョブを投入し WaitIO が通信を確立する。その際、先に投入したジョブで割り当てられた通信環境を後で投入するジョブスクリプト内で明示的に指定する必要があった。この設定を自動的に行うためのソフトウェア waitio-serv が開発されたため、ジョブ実行環境でも waitio-serv を利用するようにした。これによりジョブ投入時に後続のジョブスクリプトを毎回修正する必要がなくなった。

6. 進捗状況の自己評価と今後の展望

本年は3年計画の一年目であり、申請時に設定した表1（申請書を参照のこと）に従って、ア.不確実性定量化アルゴリズムの高度化、イ.アルゴリズムの WRF でのテスト、ウ.台風データセット version 1 の構築と性能評価、エ.データセットの解析手法の開発、という4項目に取り組んだ。

ア、イに関しては極めて順調に進捗した。すでに必要な理論立てはすべて整い、あとは実際の極端気象予測問題に適用するのみである。WRF でのテストもおおむね完了している。また今後 Wisteria/BDEC-01 で大規模問題を高速に解くための統合ソフトウェア開発も順調に進展している。

ウに関しても計画通り進展した。しかしすでに開発済みであった version 1 構築のためのコード群の Wisteria/BDEC-01 への移植に若干手間取り、データセット構築開始が想定より遅れた。この点が原因で計算資源を十分に効率よく利用することができていなかった。2年目以降、作業を加速させる必要があるものの、すでにデータ構築のためのコード群・データ群の整備はおおむね終わっているため、十分に後れを取り戻すことができると考えている。

エに関しては既に論文化した成果 (<https://www.authorea.com/doi/full/10.22541/essoar.167979634.48767251/v1>) もあり順調に進展しているが、ウ.の遅れにやや引きずられている。今後データ生成が加速すれば、さらなる成果の創出が期待できる。

以上のような進捗状況を踏まえ、今後もおおむね当初計画通り研究を進めていく。これまで開発した気象シミュレーションの不確実性定量化手法に対して統合ソフトウェアである h3-Open-SYS/WaitIO、h3-Open-UTIL/MP を適用し、Wisteria/BDEC-01 上での「計算・データ・学習」融合の実現を目指す。h3-Open-SYS/WaitIO、h3-Open-UTIL/MP の性能向上に向けた研究開発を進める他、本研究における実問題への適用事例を通じたソフトウェア改良を必要に応じて実施する。

並行してこのようなソフトウェア基盤によって

創出される台風データセットの構築とその解析によって極端気象予測の不確実性の起源の解明に迫る研究が実施できる見込みである。

7. 研究業績

(1) 学術論文（査読あり）

[1] Y. Sawada, An efficient estimation of time-varying parameters of dynamic models by combining offline batch optimization and online data assimilation, *Journal of Advances in Modeling Earth Systems*, 14, e2021MS002882, 2022

(2) 国際会議プロシーディングス（査読あり）

[2] S. Sumimoto, T. Arakawa, Y. Sakaguchi, H. Matsuba, H. Yashiro, T. Hanawa, K. Nakajima, A System-Wide Communication to Couple Multiple MPI Programs for Heterogeneous Computing, PDCAT2022 (**The Best Paper Award, Software Systems and Technologies Track**)

(3) 国際会議発表（査読なし）

[3] Nakajima, K., h3-Open-BDEC: Innovative Software Infrastructure for Scientific Computing in the Exascale Era by Integrations of (Simulation+Data+Learning), ATAT in HPSC 2023, National Central University, Taoyuan, Taiwan, March 24-25, 2023 (**Invited Talk**)

[4] Takashi Arakawa, Shinji Sumimoto, Hisashi Yashiro, Kengo Nakajima, Development of a Heterogeneous Coupling Library h3-Open-UTIL/MP, HPC challenges for new extreme scale applications, Paris, March, 2023

[5] Shinji Sumimoto, Takashi Arakawa, Yoshio Sakaguchi, Hiroya Matsuba, Hisashi Yashiro, Toshihiro Hanawa, Kengo Nakajima, Heterogeneous system for exascale using h3-Open-SYS/WaitIO, HPC challenges for new extreme scale applications, Paris, March, 2023

[6] Kengo Nakajima, Innovative Supercomputing by Integration of Simulation/Data/Learning, Paris, March, 2023

[7] Kengo Nakajima, h3-Open-BDEC: Innovative

Software Infrastructure for Scientific Computing in the Exascale Era by Integrations of (Simulation + Data + Learning), MS25: Progress and Challenges in Extreme Scale Computing and Big Data, 2023 SIAM Conference on Computational Science & Engineering (CSE2023), Amsterdam, Netherlands, Feb.26-Mar.3, 2023

[8] K. Nakajima, Innovative Scientific Computing by Integration of (Simulation + Data + Learning) in Information Technology Center, The University of Tokyo, NHR PerfLab Seminar, 2022 (**Invited Talk**)

[9] K. Nakajima, T. Iwashita, H. Yashiro, H. Nagao, T. Shimokawabe, H. Matsuba, T. Ogita, T. Katagiri, h3-Open-BDEC: Innovative Software Infrastructure for Scientific Computing in the Exascale Era by Integrations of (Simulation + Data + Learning), The 31st International Toki Conference on Plasma and Fusion Research (ITC31), 2022 (**Invited Talk**)

[10] Takashi Arakawa, Shinji Sumimoto, Hisashi Yashiro, Kengo Nakajima, Challenges of Heterogeneous Coupling, WCCM-APCOM 2022, Yokohama, 2022

[11] Kengo Nakajima, h3-Open-BDEC: Innovative Software Infrastructure for Scientific Computing in the Exascale Era by Integrations of (Simulation + Data + Learning), WCCM-APCOM 2022, Yokohama, 2022 (**Keynote Talk**)

[12] Y. Sawada, An efficient estimation of time-varying parameters of dynamic models by combining offline batch optimization and online data assimilation, American Geophysical Union Fall Meeting, 2022

[13] L. Duc, Y. Sawada, Approximation of Gradient Flows in Stein Variational Gradient Descent using Surrogate Models, American Geophysical Union Fall Meeting, 2022

[14] L. Duc, Y. Sawada, Robust ensemble Kalman filters with the Huber norm by majorization, American Geophysical Union Fall Meeting, 2022

[15] Md. R. Islam, L. Duc, Y. Sawada, Harnessing

Large Ensemble Forecast to Support Decision-Making for Storm Surge Risk Management, American Geophysical Union Fall Meeting, 2022

(4) 国内会議発表 (査読なし)

[16] 住元真司, 荒川隆, 坂口吉生, 松葉浩也, 八代尚, 埜敏博, 中島研吾, Wisteria/BDEC-01 における異種システム間連成計算実行環境, 情報処理学会研究報告 2022-HPC-185 (21), 2022

[17] 中島研吾, 住元真司, 八代尚, 荒川隆, 松葉浩也, h3-Open-BDEC: 「計算・データ・学習」融合による革新的スーパーコンピューティング, RIMS 共同研究: 数値解析が拓く次世代情報社会へエッジから富岳まで~, 2022 (**Invited Talk**)

[18] 住元真司, 荒川隆, 坂口吉生, 松葉浩也, 八代尚, 埜敏博, 中島研吾, WaitIO-Hybrid: 共有ファイルシステムと Socket を併用可能なシステム間通信ライブラリ, 情報処理学会研究報告 2022-HPC-187 (6), 2022

[19] 中島研吾, 「計算・データ・学習」融合を推進する革新的基盤ソフトウェア「h3-Open-BDEC」, 第 14 回 自動チューニング技術の現状と応用に関するシンポジウム (ATTA2022), 2022 年 12 月 23 日

[20] Takashi Arakawa, Hisashi Yashiro, Kengo Nakajima, Development of a coupler h3-Open-UTIL/MP and application, Japan Geoscience Union Meeting 2022, Makuhari, May 23, 2022

[21] 富澤風翔, 澤田洋平, Le Duc, 不確実性定量化を通じた新しい気象モデリングの可能性, 気象学会秋季大会, 2022

(5) 公開したライブラリなど

(6) その他 (特許、プレスリリース、著書等)