

全球kmスケールモデルの共通解析基盤の確立

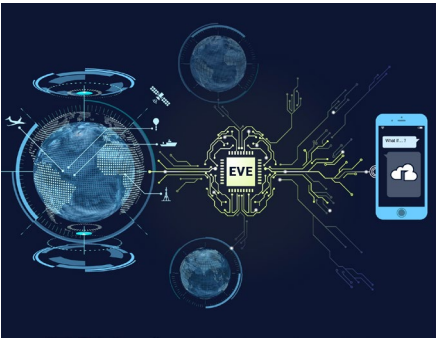
佐藤正樹^{1,2} (課題代表者) Woosub Roh¹ 大野知紀¹ 宮川知己¹ 松岸修平¹ 高須賀大輔³ 八代尚⁴ 小玉知央⁵ 荒川隆⁶

(1: Atmosphere and Ocean Research Institute, The University of Tokyo, 2: Typhoon Science and Technology Research Center, Yokohama National University, 3: Department of Geophysics, Graduate School of Science, Tohoku University; 4: National Institute of Environmental Science, 5: JAMSTEC; 6: CliMTEch Inc.)

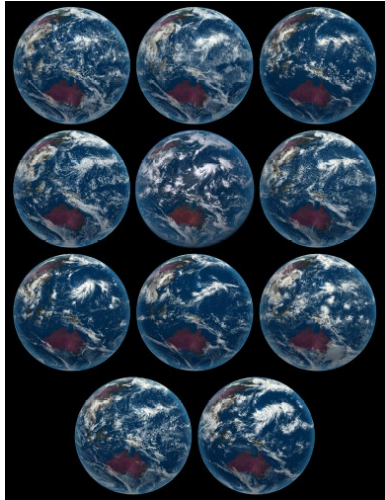
研究の目的・意義：地球全体をkmスケールのメッシュで覆う超高解像度の気象海洋モデル・地球システムモデルによる気象気候研究の新たな展開が始まっている。進行する地球温暖化に伴って激甚化している台風や豪雨等の極端気象現象のリアルタイムモニタリングや将来変化予測において、地球デジタルツインへの貢献等、全球kmスケールモデルに期待される役割は大きい。全球kmスケールモデルによるシミュレーションの性能評価のためには、世界各国の全球kmスケールモデルの比較実験を推進し、そのデータをアーカイブし、解析利用に資する必要がある。本課題では、この目的のためにJHPCNのmdxを利用する。

背景

- 世界各国の10機関以上で「全球kmスケールモデル」の開発を進めている。本課題は、世界各国の全球kmスケールモデルの出力結果をJHPCN mdxに集積し、世界中の関係する研究コミュニティへの利用を図るものである。
- 日本における「全球kmスケールモデル」は東京大学大気海洋研究所、東北大学、海洋研究開発機構、国立環境研究所、理化学研究所等の研究者の連携によって、大気海洋モデルNICAMおよびNICOCOを共同開発してきた。これらの数値モデルは、構成拠点のスーパーコンピュータ「富岳」や東京大学情報基盤センターWisteria等を用いて実施している。これらの構成拠点において実施した数値シミュレーション結果を、mdxを利用して公開することが必要とされる。その上に、世界各国で実施される「全球kmスケールモデル」の数値シミュレーション結果をmdx上にアーカイブし、世界各国の研究者の利便性に資することが必要である。
- 全球kmスケールモデルの出力は膨大であり、アーカイブされたサーバーからダウンロードすることなく、サーバー上でデータの解析する方法の確立が必要である。諸外国では、ドイツDKRZ、イギリスJASMINE等、先行する取り組みがある。JHPCN mdxでは、これら諸外国の研究拠点との連携を図り、世界的な全球kmスケールモデルのシームレスな利便性の向上を図る。



EV: <https://eve4climate.org/>



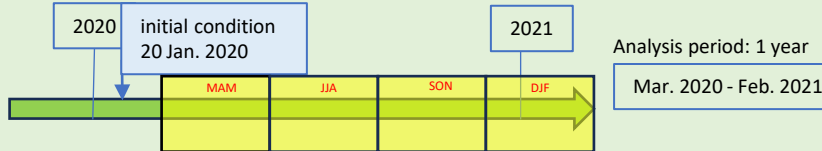
全球kmスケールモデル国際比較実験 DYAMONDの結果
Stevens et al. (2019, PEPS)

- 世界拠点 Earth Visualization Engine (EVE) の日本拠点としてデータアーカイブ解析サーバーを整備。
- 国際同時全球km-スケールモデル解析イベントHack-A-Palooza (Pan-Hackaton) を2025年5月頃に開催予定。

- 「全球kmスケールモデル」の比較実験として、1年間の積分を提唱し、世界各国の研究機関に実験の実施、データの提供を依頼する。約10機関の参加を見込む(Takasuka et al. 2024)。各モデルの0.25度のメッシュデータをmdxにアーカイブする。2024-2025年を第一フェーズの比較実験期間として、実験・データアーカイブ・解析を行い、その間の課題を明らかにし、さらに数値モデルの改良を行った上、第二フェーズの実験計画を検討する。
- mdx上では、数値シミュレーション結果の解析作業を行うことを想定している。解析作業には、現実データとの比較による数値モデルの評価・改良、地球科学的な観点からのメカニズム解明を目的とする研究の他、膨大な情報量を有するビッグデータとして情報科学的な解析の実施を想定する。
- 情報科学的な研究者と連携し、EVE (Earth Virtualization Engines; <https://eve4climate.org/>) で提唱された、AI on Top 等の構想の実現を図る。

「全球kmスケールモデル」1年比較実験プロトコル：Takasuka et al. (2024)

- Analysis period: From March 1, 2020 to February 28, 2021; Initial date: January 20, 2020
- Horizontal resolution: less than 5 km; Vertical resolution: not specified
- Physics schemes: not specified, but sub-grid convective parameterization must be switched off
- Model: Atmosphere-only model or coupled; Ensemble size: at least 1, or any number
- Output variables: 0.25 deg in the 3-hour interval for three-dimensional variables and 1-hour interval for two-dimensional variables; see Tables 4 and 5
- Initialization of land and ocean: arbitrarily up to each group



本年度の目標と進捗：

- 初年度は、国際比較実験「全球kmスケールモデル」のデータアーカイブ、配信サイトの構築を行う。10機関による全球kmスケールの計算結果を格納する予定であり、総容量はおよそ200TBとなる見込みである。mdxの大規模ストレージから200TBを課題用に確保した。
- mdx上にFTPサーバとして仮想マシンを構築した。基本機能としてのFTPサーバの構築を当初目標とした。さらに、サーバサイド解析やデータ提供フォーマットの検討を行う。仮想マシンの用途はファイルサーバでありストレージが重視される。
- 構築した仮想マシンに大規模ストレージをマウントし、FTPサーバの使用領域に設定した。またファイルのアップデートは指定されたユーザが、ダウンロードは不特定のユーザが実行できるようにFTPサーバソフトウェアの設定およびmdxのネットワーク設定を行った。これにより、本研究の第一ステップであるデータサーバ構築に関する最低限の要求を満たす。
- 複数の研究機関が個別にアーカイブした総量200TBにおよぶデータ群に対して、利用者が目的とするデータに迅速にアクセスするには、適切なデータのタグ付けとタグに基づくユーザナビゲーションの仕組みが不可欠である。今後はこれらの設計と構築についても検討する。

仮想マシンのスペック

OS	Ubuntu 22.04.3
CPU数	8
メモリ	12.09 Gbyte
仮想ディスク	100 Gbyte
高速ストレージ	20,000 Gbyte
大容量ストレージ	200,000 Gbyte

今後の展開：従来のデータ解析の手順はアーカイブされたデータをダウンロードし、ローカルなマシンで計算を行うというものである。全球kmスケールのデータは容量が大きく、アーカイブされたデータを逐一ダウンロードすることはネットワークトラフィックやディスクスペースを著しく消費することになる。そこで、必要なデータの切り出しや簡単な統計計算をサーバサイドで実行するための仕組みをmdx上に構築する。また異なる格子系で計算された結果を統一的に取り扱うためのデータ構造やソフトウェアについても設計/構築を行う。

Reference: Daisuke Takasuka, Masaki Satoh, Tomoki Miyakawa, Chihiro Kodama, Daniel Klocke, Bjorn Stevens, Pier Luigi Vidale, Christopher R. Terai (2024) A protocol and analysis of year-long simulations of global storm-resolving models and beyond. submitted to Progress in Earth and Planetary Science (submitted). <https://doi.org/10.21203/rs.3.rs-4458164/v1>



第16回JHPCNシンポジウム <https://jhpcn-kyoten.its.u-tokyo.ac.jp/ja/sympo/16th>

2024年7月11日 (木)、12日 (金) 東京コンファレンスセンター・品川

This work was supported by JAXA/EORC, JSPS Core-to-Core Program (grant number: JPJSCCA20220001). This research used computational resources of the supercomputer Fugaku (ID: hp220058, hp220132, hp230078, hp230278, hp240106)