

令和 5 年 6 月 26 日現在

機関番号：82101

研究種目：挑戦的研究（萌芽）

研究期間：2019～2022

課題番号：19K22853

研究課題名（和文）時間方向並列化と連成カプラーを用いた超高解像度・長期気候シミュレーションの革新

研究課題名（英文）Innovation of high-resolution long-term climate simulation using parallel-in-time method and coupling library

研究代表者

八代 尚（Yashiro, Hisashi）

国立研究開発法人国立環境研究所・地球システム領域・主任研究員

研究者番号：80451508

交付決定額（研究期間全体）：（直接経費） 4,100,000円

研究成果の概要（和文）：より精度の高い気候変動予測を実現することが期待される、高解像度気候シミュレーションの計算速度限界を打破するために、時間方向並列化の手法を用いた階層型気候計算基盤を構築することを目的に研究を行った。異なる2つの空間解像度を持つ大気モデルを連成実行するライブラリを活用し、時間方向並列化のための技術開発を行い、その有効性について明らかにした。研究の結果、より多数の計算機を用いた十分な高速化のためには、低解像度シミュレーションの高速化と高精度化が鍵であることを示し、その実現方法として高解像度計算結果と機械学習手法を用いたデータ駆動型モデルコンポーネントの構築について検討を行った。

研究成果の学術的意義や社会的意義

より正確な将来の気候変動予測を行うことは喫緊の課題であり、気候モデルシミュレーションの不確実性の低減のためにはモデルの時空間的な高解像度化が必要不可欠である。特に粗視化による仮定を多く置いている全球50-100kmメッシュから10kmメッシュ以下での計算に移行することが重要である。一方で、近年のスーパーコンピュータは並列度を増加させることで性能と電力効率の両立を図っており、気候モデルが現状の並列化方法のみでそれに対応している、今後どれだけ大規模な計算機が登場したとしても、高解像度気候計算の高速化は困難である。本研究はこの気候の計算科学が直面している一般性の高い問題を解決するものである。

研究成果の概要（英文）：To realize more accurate prediction of future climate change, it is required to make a breakthrough about the computational speed-up of high-resolution climate simulation. This study aims to develop the hierarchical climate simulation infrastructure by using the parallel-in-time (PinT) method. We used the library to execute coupled simulations between the models with different horizontal resolution and evaluated the effectiveness of the PinT method. We found that the further improvement of accuracy and speedup of the low-resolution climate simulation is necessary for the better scalability of PinT computation. For this purpose, we studied about the development of data-driven model components using machine-learning method using high-resolution simulation results.

研究分野：大気科学

キーワード：高性能計算 気候変動予測 連成計算 並列計算 モデルシミュレーション

1. 研究開始当初の背景

気候研究において、シミュレーションモデルの空間解像度は重要な課題である。水平解像度が 0(100km)の全球大気循環モデルでは、格子間隔よりも粗い気象現象、特に積乱雲のような対流圏内で大きく発達する対流雲を陽に表現することが出来ない。そのため現象を統計的に扱ったり、簡素化したパラメタリゼーションモデルが導入したりされている。対流パラメタリゼーション(CP)モデルの不確実性が現在気候の再現性や将来予測性能に与える影響は大きく、この不確実性を低減することが重要な課題となっている。この問題の一つの解決策として、CPモデルを用いず、空間解像度を高めることで対流雲を陽に計算し、結果を改善しようという試みが進められている。Kodama et al.(2015)は、スーパーコンピュータ「京」を用いて 14km メッシュで 20 年という長期シミュレーションを行い、これまでの気候モデルでは得られなかったような現在気候の卓越した再現性を示した。一方で、近年のスーパーコンピュータ(スパコン)システムの変遷は、気候予測計算において必ずしも歓迎されるものではない。気候シミュレーションは $O(10^8)$ を越える時間ステップを必要とし、並列コンピュータを用いて高速化を図る場合、ストロングスケールリング問題となる。図 1 は全球非静力大気モデル NICAM と東京大学のスパコン Oakforest-PACS を用いてシミュレーションを行った際のストロングスケールリング性能結果である。黒線は 224km メッシュでの計算結果であり、多数の MPI プロセスを用いると並列度が不足し、システム全体の 1%に満たないノード数で既に計算速度の向上が頭打ちになることがわかる。灰線は 56km メッシュでの計算結果であり、224km よりも格子数は 16 倍、時間刻み幅が 1/4 であるため時間ステップは 4 倍増加している。プロセス分割数を増やしてより高速化しようとする、性能の頭打ちは低解像度の時よりも早く訪れる。これは時間ステップの増加分をより多く空間方向にプロセス分割することで補おうとするためである。この結果は、どれだけ多くのノード数を持ったスパコンがあっても、現状の空間方向並列化のみに頼っている場合、高解像度気候計算はストロングスケールリング性能の限界により、現行の低解像度計算よりも高速化することができないというジレンマを我々に提示するものである。これは多くの気候モデルに共通する課題である。

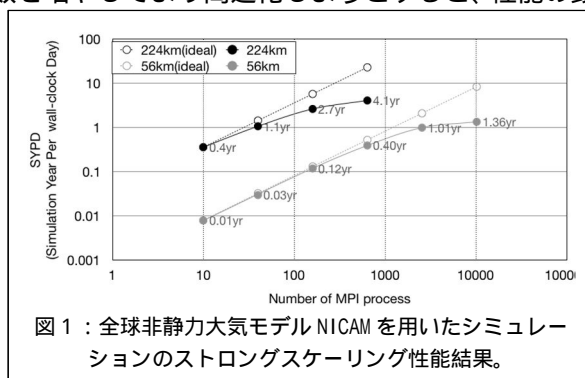


図 1 : 全球非静力大気モデル NICAM を用いたシミュレーションのストロングスケールリング性能結果。

2. 研究の目的

本研究は、時間方向並列化の手法を用いることにより、要求される並列度が年々増大しているスーパーコンピュータの性能を十分に引き出し、数十年から数百年にわたる気候シミュレーションを従来の 10 倍以上の水平解像度で、かつより高速に計算可能な階層型シミュレーション基盤を構築することを目的とする。しかし、複雑な諸物理過程を多数有し非線形性の強い大気シミュレーションの計算結果を、粗い計算と詳細な計算の間で高精度・低反復回数を保って収束させることは困難である。そこで気候シミュレーションで収束が要求される要素に限定し判定基準を探索することで、実用に耐え得る高速化を達成することを目指す。

3. 研究の方法

雲をモデル内で陽に表現し将来予測精度の不確実性を減らすことの出来る高解像度モデル(雲解像モデル)を長期気候シミュレーションで活用するために、Parareal 法²⁾を用いた時間方向並列化を大気モデルに適用することを試みる。Parareal 法では高速で粗視化されたシミュレーションと、計算時間がかかる詳細なシミュレーションを組み合わせ、粗視化シミュレーションの結果を修正しながら反復計算し、結果を収束させる。目的でも述べたとおり、非線形性の強いシミュレーションにおいて、雲の分布を含めた大気の状態を高精度・低反復回数で収束させるのは困難である。そこで本研究では、気候学的観点に立った比較的緩い判定基準での高速な収束計算を目指す。すなわち、気候シミュレーションにおいて時間的連続性が保たれてほしい地表面気温や大気の放射収支などに着目した収束基準を選定し実験を行う。

空間解像度の異なる 2 つの計算を連成させるシステムには、気候モデルで海洋モデルと大気モデルを結合する用途等に開発されたカップリングライブラリを拡張し、データの受け渡しと

空間解像度の変更を担当させる。これにより、数十万行に渡る巨大な大気モデルのソースコードをほとんど変更することなく、Parareal 法を適用することが可能になる。低解像度シミュレーションの修正計算方法には、高解像度シミュレーションの結果をリファレンスとして計算した強制項をある時定数で緩和させながらシミュレーションを進めるナッジング法の適用を試みる他、高解像度と低解像度でそれぞれ計算される計算結果の違いを、新たな物理過程のひとつとして追加する方法を試みる。

4. 研究成果

<連成システムの構築>

まず、時間方向並列化手法を気候シミュレーションに適用するための準備として、空間解像度の異なる 2 つの大気シミュレーションを連成させるシステムの構築を行なった。大気モデルには全球非静力正 20 面体大気モデル NICAM (Satoh et al., 2014) を用いた。一般的に気候モデルは、大気の流体力学を解く部分について空間解像度に依らないスキームを採用している。一方で、諸物理過程はその定式化、パラメタリゼーション方法がある時空間解像度を想定して行われているため、 $O(100\text{km})$ スケールの低解像度モデルと $O(1\sim 10\text{km})$ スケールの高解像度モデルでは用いられるスキームが異なる。そのため、低解像度と高解像度の両方に NICAM を用いるが、それぞれのシミュレーション結果の挙動は異なる。2 つの大気シミュレーションはそれぞれが並列計算を行っており、これをカップリングライブラリ Jcup (Arakawa et al., 2014, 2020) で結合しながら並行して実行する。Jcup は汎用ライブラリであり、気象・気候シミュレーションに限らず複数の独立したアプリケーション間で変数を交換することが可能である。また、それぞれのシミュレーションモデルが異なる格子系を用いている場合、お互いのモデル格子点に合うように空間方向の値の再配分 (リマッピング) を行うことが可能である。本研究では、低解像度と高解像度の NICAM 格子点の間で値を交換するための、マッピングテーブル作成機能を開発した。NICAM は流体力学計算において離散化ノイズを抑える工夫として、ばね格子を用いた格子点の再配置を行なっている。そのため、低解像度格子点の配置と高解像度のそれとは一致しない。そのため各格子が代表するコントロール面積の配分アルゴリズムについては、Jcup がもともと有している汎用的な方法を採用して開発を進めた。

図 2 に、低解像度と高解像度のシミュレーション間のデータ受け渡しの概略を示す。ここからは、高解像度シミュレーションを HiRes、低解像度シミュレーションを LoRes と表記する。HiRes で計算された大気の状態変数 (風速、気温等) は Jcup を介してリマッピングされながら LoRes に受け渡される。このリマッピングにより、HiRes の計算結果は粗視化され、大気の状態を空間方向に波数スペクトル分解した際の低波数成分のみが対象となる。次に、粗視化された HiRes と LoRes の差を元に、ある時定数での修正項 (テンデンシ) が計算される。これは LoRes の結果をリファレンスとしたニュートン緩和 (ナッジング) と同様の意味を持つ。計算されたテンデンシは再び Jcup を通じて HiRes へ受け渡され、各格子点での状態量の時間発展が解かれる。このようなデータ交換を行うことで、HiRes が再現する細かな空間スケールでの変動を弱めることなく、LoRes によって計算されたより空間スケールの大きな変動のみを作用させたシミュレーションが可能になる。逆に、HiRes の結果を LoRes に作用させる場合は、LoRes に送られたデータを用いて LoRes の時間発展を解くのみである。開発した連成計算システムを用いて試験計算を実施したところ、2 つの解像度で用いられている標高データの値の違いに起因するエラーが結果の悪化を引き起こすことが判明した。NICAM は地形に沿った鉛直座標系を採用しており、標高が異なる場合、同じ鉛直層番号であっても絶対高度が異なる。この問題を解決するため、Jcup を介した標高情報のモデル間交換を追加し、LoRes 側でテンデンシを計算する際に高度補正処理を行うことにした。これにより、標高の不整合によるエラーを除去することに成功した。加えて、データ交換時の補間方法についても追加検討を行なった。これまで Jcup が扱ってきた事例は大気・海洋モデル間の熱・水交換など、エネルギー収支や物質収支が閉じなければいけないタイプのデータ交換であったため、デフォルト

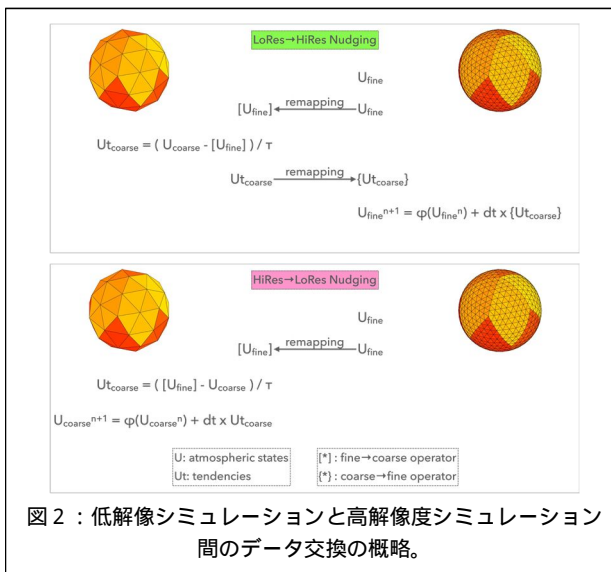
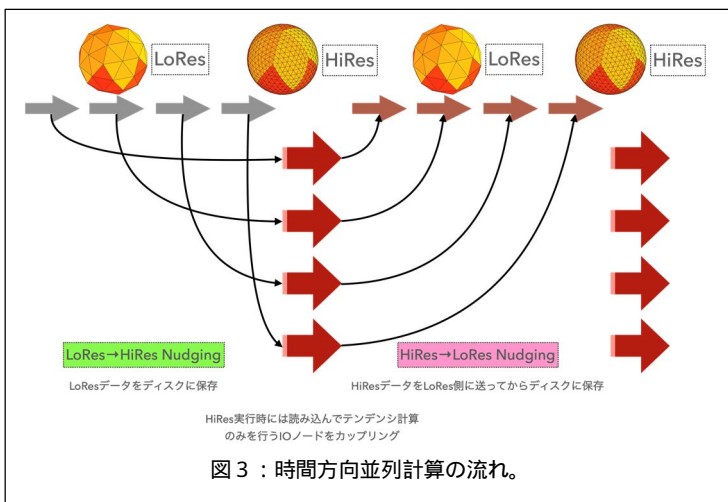


図 2: 低解像シミュレーションと高解像度シミュレーション間のデータ交換の概略。

の補間方法として面積重み付き内挿を採用してきた。しかし、この方法では LoRes から HiRes へデータを転送した際に、与えられたテンデンシが LoRes の格子形状のまま階段状に入力されてしまうことが実験を進める中で判明した。そこで新たに線形内挿補間方式を選択できるようカップリングライブラリの機能を拡張し、LoRes から HiRes へ受け渡すテンデンシの補間については線形内挿で与えられるようにした。これにより、本来意図していた HiRes 側での低波数成分飲み状態量補正をより正しく行うことが可能になった。

<時間方向並列計算の演算性能評価>

連成計算システムの開発によって実施可能になった時間方向並列計算について、実験デザインと計算性能評価を行なった。時間方向並列化のためのシステムでは、ひとつの LoRes の結果を用いて開始時刻の異なる複数の HiRes が同時実行される。また、時間方向に分割された HiRes の結果をすべて用いて次の反復ステージにおける LoRes の補正を行う。これらをリアルタイムの連成計算のみで実現しようとした場合、メモリ上に一時的に確保しておくデータ量が増加してしまうため、実験そのものを行うシステムの制限が厳しくなってしまうことがわかった。そこで、オンラインのみで完結することにこだわらず計算結果を適宜ファイルに書き出し、利用メモリ量を抑えることとした。図3に時間方向並列計算の模式図を示す。本来の Parareal 法では、詳細度の異なる2つのシミュレーション間の情報のやり取りはリスタート時刻での状態量の一括修正を行うが、本研究では大気シミュレーションが安定して実行されることを重視し、お互いに相手の状態量を用いて修正を行う際に、ゆっくりと緩和させることとした。この時の時定数は、トレーサー輸送実験等で大気再解析データを用いてナッジングを行う場合に倣い、風速を1日、気温を5日に設定した。また、各 HiRes 計算ではスピナップ期間としてターゲット年の開始の日の1か月前から計算を行うこととした。HiRes を実行する際には、LoRes の計算結果を読み込んでテンデンシを計算し、また LoRes の解像度にマッピングされた HiRes の結果をファイル出力する10ノードをカップリングして実行することで、オンライン計算と同様の連成を実現し、また HiRes の結果を出力する際のデータサイズを抑えることに成功した。



このオフライン併用の時間方向並列計算について、東大 Oakbridge-CX スーパーコンピュータ (OBCX) を用いて実験を実施し、OBCX を用いた場合の 224km-56km 連成と 224km-14km 連成の 10 年分の計算時間について実測と机上での見積もり計算を行った結果を図4に示す。まず、56km メッシュ計算(図4上)では、通常の高解像度計算のみを利用ノード数を増やして高速化しようすると、ノード数の増加に伴って10年分の気候実験(大気シミュレーションのみ)にかかる計算時間が減少していくが、計算時間の減少率は徐々に悪化し、256ノード利用では64ノード利用と比較してもう高速化されなくなることがわかる。これに対して224km-56km 連成での時間並列計算では、さらに多数の計算ノード(640ノード)を用いて10年分の高解像度計算を1年ずつ並列に実行することで、高解像度計算にかかる時間を削減することに成功している。しかし、224km メッシュの計算を行うための時間が余計にかかり、また時間方向並列計算を収束させるための反復回数によっては、利用ノード数を増加させただけの効率的な高速化は見込めないことがわかった。この実験設定では、64ノードを用いたこれまでの計算と、640ノードを用いて反復回数3回を仮定した時間方向並列計算で、それぞれ1.1SYPD (Simulation Year Per wallclock-Day, 実時間1日で計算可能な年数)と1.0SYPD という値を得た。一方14km メッシュのケース(図4下)では、従来のシーケンシャルな計算手法で1024ノードまで利用ノード数を拡大した場合の計算時間は0.4SYPDと見積もられた。14km メッシュ解像度においてもやはりノードスケール(ストロングスケール)による計算高速化は1024ノードあたりでほぼ頭打ちとなっている。これに対し、224km-14km 連成での計算では仮に OBCX 総ノード数(1368ノード)を越える10240ノードが利用可能であれば、反復回数3~5回を仮定した場合0.7-0.4SYPDで計算可能であると見積もられた。この空間解像度であれば、高速化率は低い、ストロングスケールでこれ以上高速化不能であった壁を、時間並列化によって打破できることが示された。

<低解像度シミュレーションの精度向上に向けたパラメタリゼーションモデル構築>

ここまでの研究成果によって、時間方向並列化の高速化のためには、収束計算のための反復回数を少なく済ませることで、LoRes の計算にかかる時間をより短縮させることが必要であることがわかった。反復回数が増加する要因として、HiRes と LoRes の互いのシミュレーションに現れる気候バイアスの傾向が異なることが挙げられる。気候バイアスは解像度によって用いる物理コンポーネントスキームが異なることに起因している。そこで、LoRes のバイアス傾向を HiRes に近づけ、またより高速に計算を行うことを目的として、HiRes で計算された特定のモデルコンポーネントの入出力結果を活用し、AI 技術を用いたデータ駆動型モデルを構築することを試みた。このようなデータ駆動型のパラメタリゼーションモデルは代理モデル (Surrogate Model) と呼ばれる。本研究では諸物理課程のうち、特に解像度によるバイアス傾向が大きく異なる、雲モデルを対象とした。学習のための教師データを効率よく準備するために、他の研究課題[中島ら、科研費基盤 S]にて開発された h3-Open-UTIL/MP ライブラリを用いた。h3-Open-UTIL/MP は Jcup を拡張したカップリングライブラリであり、Fortran で記述された HPC アプリケーションと Python アプリケーションを連成した計算を行うことが可能である。Jcup の機能を引き継ぐことで、シミュレーションモデルの任意の位置から受け渡す変数を抽出し、同時にリマッピングによって解像度を変換することが可能となっており、HiRes 側で得られた大容量の途中計算結果を、効率よく LoRes 側の機械学習ツールに供給し、粗視化された代理モデルを構築する上で有効である。このライブラリを用いて、14km メッシュ計算で用いられるシングルモーメント雲微物理スキームの粗視化代理モデルの構築を試みた。機械学習ツールには PyTorch を使い、各時刻ステップの中で雲微物理スキームに入力された変数と、出力として得られた気温、密度、各水物質の質量存在比の変化率のデータペアを教師データとして供給した。構築されたデータ駆動型モデルコンポーネントは LoRes シミュレーションの計算時間を 20-30%削減すると評価された。また、データ駆動モデルの予測精度をさらに向上するための手法として、HiRes の計算結果を LoRes のメッシュサイズに粗視化する前にもともと持っていた情報を統計的情報あるいは畳み込みニューラルネットワークの入力として活用する方法について検討し、それらの情報をリマッピング (= 平均値の導出) とは別に LoRes 側の機械学習に受け渡すためのライブラリの機能拡張を行った。データ駆動型の粗視化代理モデルは機械学習のネットワークモデルの選択やチューニングの余地があり、今後の研究課題として更なる取り組みが必要である。

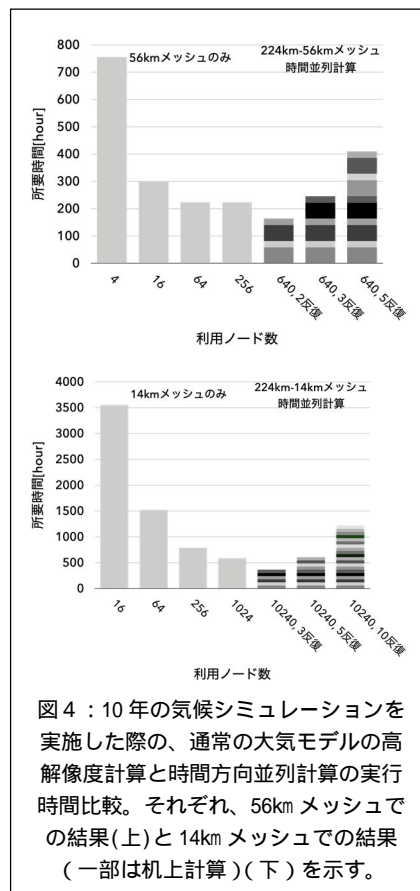


図4：10年の気候シミュレーションを実施した際の、通常の大気モデルの高解像度計算と時間方向並列計算の実行時間比較。それぞれ、56kmメッシュでの結果(上)と14kmメッシュでの結果(一部は机上計算)(下)を示す。

参考文献

1. Kodama, C. et al., (2015): A 20-year climatology of a NICAM AMIP-type simulation. *Journal of the Meteorological Society of Japan*, **93**, 393-424. <http://doi.org/10.2151/jmsj.2015-024>.
2. Satoh, M., H. Tomita, H. Yashiro, H. Miura, C. Kodama, T. Seiki, A. T. Noda, Y. Yamada, D. Goto, M. Sawada, T. Miyoshi, Y. Niwa, M. Hara, Y. Ohno, S. Iga, T. Arakawa, T. Inoue, and H. Kubokawa (2014): The Non-hydrostatic Icosahedral Atmospheric Model: Description and Development. *Progress in Earth and Planetary Science*, **1**, 18. doi:10.1186/s40645-014-0018-1
3. Lions, J-L, Y. Maday, G. Turinici (2001): A "parareal" in time discretization of PDE's, *Comptes Rendus de l'Académie des Sciences, Série I*. Elsevier. **332** (7), 661-668, doi:10.1016/S0764-4442(00)01793-6
4. Arakawa, T., T. Inoue, and M. Sato (2014): Performance evaluation and case study of a coupling software ppOpen-MATH/MP. *Procedia Comput. Sci.*, **29**:924-935
5. Arakawa, T., Inoue, T., Yashiro, H. et al. Coupling library Jcup3: its philosophy and application. *Prog Earth Planet Sci* **7**, 6 (2020). <https://doi.org/10.1186/s40645-019-0320-z>
6. 科研費基盤研究S,(計算+データ+学習)融合によるエクサスケール時代の革新的シミュレーション手法(研究代表者:中島研吾、2019~2023年度)

5. 主な発表論文等

〔雑誌論文〕 計1件（うち査読付論文 1件／うち国際共著 0件／うちオープンアクセス 1件）

1. 著者名 Arakawa Takashi, Yashiro Hisashi, Nakajima Kengo	4. 巻 1
2. 論文標題 Development of a coupler h3-Open-UTIL/MP	5. 発行年 2022年
3. 雑誌名 HPCAsia2022: International Conference on High Performance Computing in Asia-Pacific Region	6. 最初と最後の頁 72-83
掲載論文のDOI（デジタルオブジェクト識別子） 10.1145/3492805.3492809	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

〔学会発表〕 計9件（うち招待講演 1件／うち国際学会 2件）

1. 発表者名 Yashiro, H.
2. 発表標題 Challenges to large-scale weather/climate simulation for disaster prevention and mitigation, The International Conference on High Performance Computing in Asia-Pacific Region
3. 学会等名 The International Conference on High Performance Computing in Asia-Pacific Region (HPC Asia 2022) (招待講演) (国際学会)
4. 発表年 2022年

1. 発表者名 Yashiro, H., T. Arakawa, S. Sumumoto, K. Nakajima
2. 発表標題 h3-Open-UTIL/MP: A General-purpose Coupling Library Bridging Legacy HPC Applications and the Future
3. 学会等名 International Workshop on the Integration of (Simulation + Data + Learning) : Towards Society 5.0 by h3-Open-BDEC (国際学会)
4. 発表年 2021年

1. 発表者名 八代尚
2. 発表標題 大規模気象・気候計算への挑戦と課題
3. 学会等名 「富岳」成果創出加速プログラム 研究交流会「富岳百景」
4. 発表年 2022年

1. 発表者名 八代尚, 寺崎康児, 中野満寿男, 小玉知央, 河合佑太, 富田浩文
2. 発表標題 計算機システムと気象・気候分野のアプリケーションのコードデザイン: NICAM-LETKFの経験から
3. 学会等名 日本気象学会2021年度秋季大会
4. 発表年 2021年

1. 発表者名 八代尚
2. 発表標題 気象・気候アプリケーションソフトにおける緩やかな変革、また異なる計算手法との緩やかな融合に向けて
3. 学会等名 PCクラスターワークショップ in 柏 2021
4. 発表年 2021年

1. 発表者名 八代尚
2. 発表標題 超重量級計算はこの先どうなるのか?: 気象・気候シミュレーション研究が期待するもの
3. 学会等名 PCクラスターシンポジウム
4. 発表年 2022年

1. 発表者名 八代尚
2. 発表標題 国内外の気象・気候シミュレーションモデル開発動向について
3. 学会等名 日本気象学会第1回計算科学研究連絡会
4. 発表年 2022年

1. 発表者名 八代尚
2. 発表標題 全球大気モデルNICAMの最近の取り組み紹介
3. 学会等名 次世代陸モデル開発・応用・社会実装に関する合同ワークショップ
4. 発表年 2021年

1. 発表者名 八代尚
2. 発表標題 気象シミュレーションモデル開発と計算機に関する近年の動向
3. 学会等名 今後の台風予測研究に関する研究会
4. 発表年 2021年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関