

[JHPCN シンポジウム]

令和 4 年度 JHPCN 採択課題ポスター紹介

令和 4 年度学際大規模情報基盤共同利用・共同研究拠における当センターとの共同研究課題のポスター5 件と Web サイト 2 件をご紹介します。

以下 5 件の課題は、次ページ以降のポスターをご覧ください。

- ・ jh220018 「日本全土の洪水氾濫被害と適応策の検討」

研究代表者 峠嘉哉(東北大学)

- ・ jh220025 「センター間連携による柔軟な計算資源提供に関する研究」

研究代表者 滝沢寛之(東北大学)

- ・ jh220027 「非同期入出力機構を用いた大規模乱流直接数値シミュレーションコードの開発」

研究代表者 横川三津夫(神戸大学)

- ・ jh220043 「熱中症リスク評価シミュレータの開発と応用」

研究代表者 平田晃正(名古屋工業大学)

- ・ jh220050 「Implementation and Application of High-Performance Empirical Dynamic Modeling」

研究代表者 高橋慧智(東北大学)

以下の 2 課題については URL をご参照ください。

- ・ jh220048 「HPC と高速通信技術の融合による大規模データの拠点間転送技術開発と実データを用いたシステム実証試験」

研究代表者 村田健史(国立研究開発法人情報通信研究機構)

<https://jhpcn-kyoten.itc.u-tokyo.ac.jp/abstract/jh220048>

- ・ jh220053 「プロペラ駆動小型無人機的设计検討技法の確立を目指した空力・推進・構造の実機丸ごと統合シミュレーション」

研究代表者 金崎雅博(東京都立大学)

<https://jhpcn-kyoten.itc.u-tokyo.ac.jp/abstract/jh220053>

学際大規模情報基盤共同利用・共同研究拠点 公募型共同研究 令和4年度採択課題

14th Symposium

jh220018

Joint Usage / Research Center for Interdisciplinary Large-scale Information Infrastructures

峠 嘉哉 (東北大学工学研究科), 滝沢 寛之 (東北大学サイバーサイエンスセンター)

日本全土の洪水氾濫被害と適応策の検討

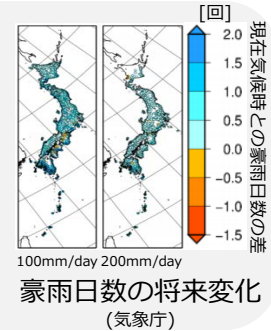
JHPCN

背景・目的

近年、日本各地で洪水被害が多発しており、
気候変動の影響で豪雨の増加が見込まれる
洪水リスク上昇地域の最適な対策は何か
⇒ 将来への効率的な治水投資
複数の適応策の比較と地域性考慮が必要



山形豪雨 (2020年)



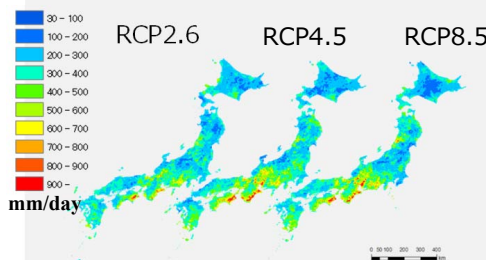
豪雨日数の将来変化 (気象庁)

**研究目的：日本全国でのシナリオ（気候や社会変化、適応策）
に応じた膨大な組合せ氾濫被害推定の高速度・高精度化**

洪水氾濫解析

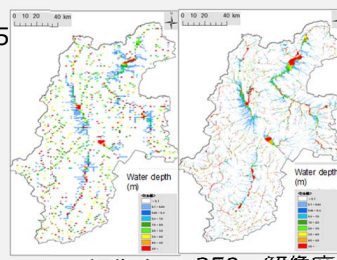
(Yamamotoら, Climatic Change, 165:60)

将来の降水量：全球気候モデル



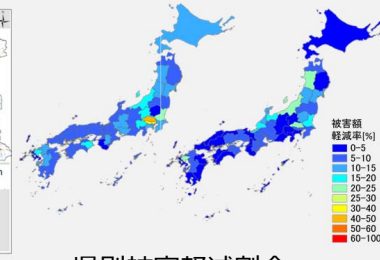
日本各域で再現期間100年洪水
を生じさせる日降水量の気候シ
ナリオRCPの違い (2100年時)

高解像度洪水氾濫解析



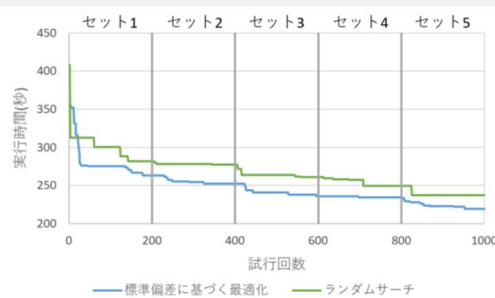
1km解像度 250m解像度
100年確率降雨時の
浸水深分布

適応策の被害軽減比較



県別被害軽減割合
左：治水施設レベルの向上
右：たんぼダム

並列化の負荷分散



各プロセスの実行時間の標準偏差を目的
関数とし、これを最小化したベイズ最適
化とランダムサーチの比較
(石塚ら, 情報処理学会論文誌, 59)

多数の組み合わせ計算

複数計算の効率化・高速化

膨大な計算量

シナリオ3
RCP2.6+SSP1
RCP4.5+SSP2
RCP8.5+SSP5

将来目標3
現在
2050年頃
2100年頃

気候モデル5
国内2
豪州1
仏国1, 独国1

リスク計算：再現期間4, 30, 50, 100, 200年

適応策5：適応なし、治水施設、土地利用など

900通り ×
全国600万セル

並列化の負荷をベ
イズ最適化を利用し、
高速化をはかる

JHPCN

学際大規模情報基盤共同利用・共同研究拠点 第14回シンポジウム

Japan High Performance Computing and Networking plus Large-scale Data Analyzing and Information Systems

2022年7月7日, 8日

HYBRID

学際大規模情報基盤共同利用・共同研究拠点公募型共同研究 令和4年度採択課題

14th Symposium

jh220025

Joint Usage / Research Center for Interdisciplinary Large-scale Information Infrastructures

滝沢寛之(東北大学) 佐々木大輔(金沢工業大学) 伊達進(大阪大学)

江川隆輔(東京電機大学) 小川泰一郎(大阪公立大学) 下村陽一(東北大学) 高橋慧智(東北大学) 高橋俊(東海大学)

センター間連携による柔軟な計算資源提供に関する研究

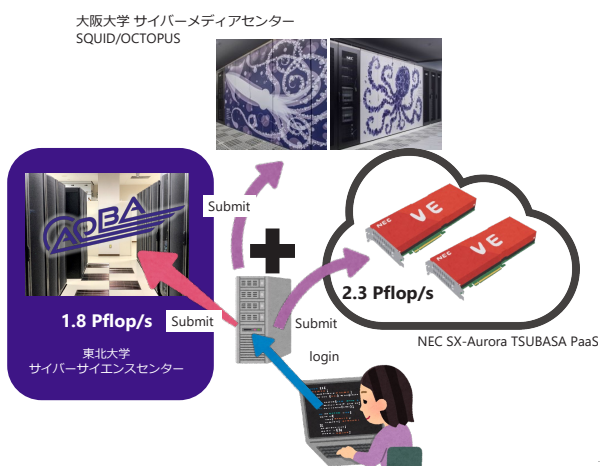
JHPCN

研究目的 JHPCNは8つの機関から構成されるネットワーク型の共同利用・共同研究拠点であり、それぞれの機関が独自の計算資源を独自の運用方針に基づいて提供している。本研究の目的は、JHPCN計算資源の有効活用を実現する資源管理機構とその運用体制を検討し、実現可能性を明らかにすることである。各計算資源の利用率は時々刻々と変化しており、ある機関で計算資源が足りない時に、他の機関では同じ種類の計算資源が遊休状態にある可能性がある。このため、より柔軟に計算資源を融通しあう環境を整備することができれば、JHPCN全体で計算資源のより効率的な活用を期待できる。そのような複数機関の連携による運用を実現するためには、ストレージシステムやバッチジョブスケジューラの連携という技術的検討に加えて、課金体系や利用者管理などの運用方針の観点からも様々な検討が必要となる。本研究では、技術的視点からの検討によって実現可能性を示すとともに、複数のセンター間連携による運用の実現に向けた運用上の課題の明確化を行う。

研究計画 東北大学サイバーサイエンスセンターと大阪大学サイバーメディアセンターとの間で相互にジョブを投入する仕組みを構築し、同センター間でのクラウドバースティングを実現する。東北大学サイバーサイエンスセンターで開発中のジョブスケジューリングのシミュレータを用いて、ジョブスケジューラの設定について検討する。近年急速に需要が高まっているデータ駆動型研究のジョブは、巨大なデータを扱うために従来の数値シミュレーションよりも他センターへ転送するためのコストが大きい。このことを踏まえ、ジョブスケジューリングのアルゴリズムや適切なパラメータ設定を議論する。以上の技術的な検討項目に加えて、例えば課金制度や利用者管理、利用者支援業務など、連携して運用する際の両センターの運用方針に関する課題を明確化する。

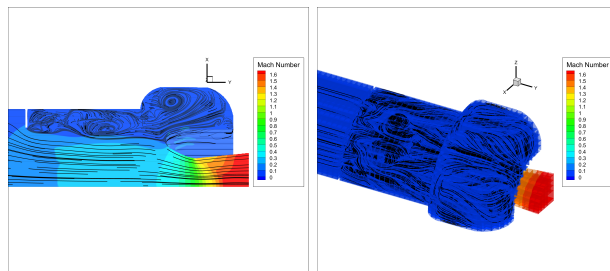
センター間ジョブ転送の仕組みの構築

本研究では、東北大学サイバーサイエンスセンターと大阪大学サイバーメディアセンターのそれぞれのセンターにおいて専用のバッチキュー(以下、転送キューと呼ぶ)を設け、そこに投入されたジョブは他方のセンターの計算資源を使って実行されるように設定する。そのようなジョブ転送の仕組みを実現することを目標とし、その実現に必要な基盤ソフトウェアの研究開発を行う。



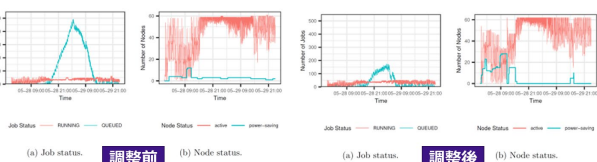
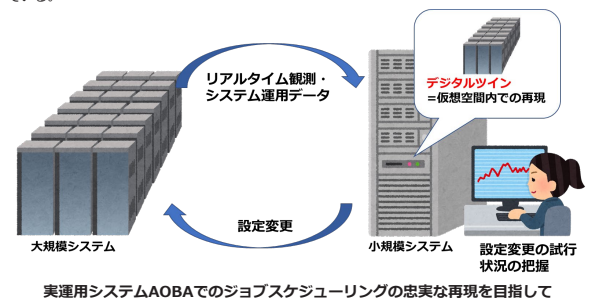
大阪大学サイバーメディアセンターではオンプレミスの計算資源からパブリッククラウドへのジョブ転送(クラウドバースティング)をすでに運用しており、東北大学サイバーサイエンスセンターでもクラウドバースティングの運用を予定している(2022年10月運用開始)。このことから、パブリッククラウドへの転送とセンター間転送との適切な使い分けも議論する。

実証実験においては、共同研究者の数値流体シミュレーション(極超音速流れの解析、移動物体(膜翼)の解析、ロケット燃焼室内3次元流れ)を題材とし、それぞれの研究分野での成果創出を旨と同時に、転送キューにジョブを投入して実アプリケーションの長時間実行を実施する。その際の使い勝手やオーバーヘッドの情報を収集し、その後の研究開発の方針を検討する際に役立てる。

キャビティ付き固体燃料ロケット内部流れ^[1]

実運用システムにおけるジョブスケジューリングのシミュレーション

東北大学サイバーサイエンスセンターでは、Slurmシミュレータに独自の拡張を加えて同センターが運用するスーパーコンピュータAOBAのジョブスケジューリングを再現する研究開発を行っている。

パラメータ調整による効果をシミュレーションによって事後検証^[2]

現在ではAOBA内でのジョブスケジューリングを対象とした研究開発を行っているが、パブリッククラウドや他のセンターへのジョブの転送も可能になった場合を想定して、その適切なジョブスケジューリングを実現するための検討を行う。

具体的には、ジョブ転送のオーバーヘッドをモデル化し、どのジョブを転送キューに移せばシステム全体としての効率を改善することができるかを予測する手法を検討する。サイト間のデータ転送のコストが大きいことを考えれば、利用するデータ容量の大きいジョブの転送を回避する方針が有望だと予想される。そのようなモデル化に基づいて、選択的にジョブを転送することでシステム全体としての効率に与える影響をシミュレーションに基づく調査により明らかにする。

運用面での課題の明確化

本研究課題で実証実験に必要な計算資源の利用負担金は本研究課題で負担することを前提に協力を募る計画になっているが、実際の運用を念頭に置いた場合には、利用者管理や課金処理において多くの課題がある。本研究課題を通じてこれらの課題を明確化するとともに、その解決に向けた運用方針についても具体的に検討する。

参考文献

- [1] 小川泰一郎, 大内健太郎, 佐々木大輔, “直交格子積み上げ法を用いたキャビティ付き固体燃料ロケット内部流れの流体解析手法の構築,” 第54回流体力学講演会/第40回航空数値シミュレーション技術シンポジウム, 1C02, 2022.
- [2] Tatsuyoshi Ohmura, Yoichi Shimomura, Ryusuke Egawa and Hiroyuki Takizawa, “Toward Building a Digital Twin of Job Scheduling and Power Management on an HPC System,” 25th Workshop on Job Scheduling Strategies for Parallel Processing (JSSPP 2022), Online, June 2022

JHPCN

学際大規模情報基盤共同利用・共同研究拠点 第14回シンポジウム

Japan High Performance Computing and Networking plus Large-scale Data Analyzing and Information Systems

2022年7月7日, 8日

東京コンファレンスセンター・品川

JHPCN: 学際大規模情報基盤共同利用・共同研究拠点 第14回シンポジウム 2022年度課題
2022年7月7日(木) @東京コンファレンスセンター・品川

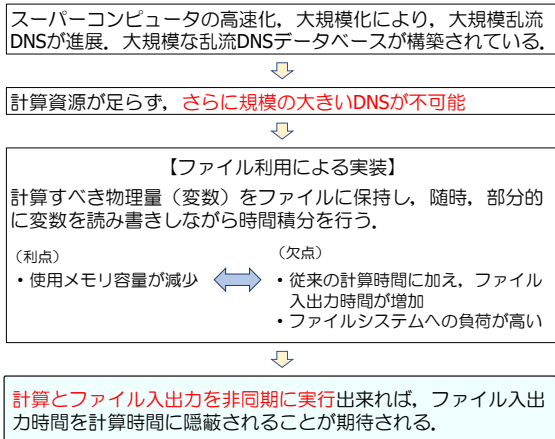
jh220027 非同期入出力機構を用いた大規模乱流直接数値シミュレーションコードの開発

代表：横川 三津夫（神戸大学）

撫佐昭裕，小松一彦（東北大学），石原卓（岡山大学），曾我隆，山口健太，山根悠輝（日本電気）
松本泰生，畑中裕翔（神戸大学）

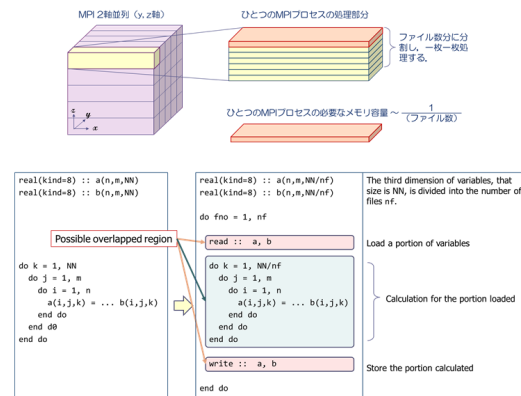
本課題では、非圧縮性及び圧縮性乱流直接数値シミュレーション（Direct Numerical Simulation, DNS）コードに対し、変数の一時退避場所としてのテンポラリ・ファイルを作成し、使用メモリ容量を削減するコード実装を行い、その計算性能及びファイル入出力性能の評価を行う。特に、ファイル入出力に対しては、SX-Aurora TSUBASA が提供する高速入出力機構や非同期入出力機構を利用し、大規模DNS実行の可能性について評価する。

1. はじめに — 背景とコード実装の考え方



3. 実装イメージ

並列化コードで使用されている変数に対し、1つのプロセス中の変数をさらに分割し、個々のファイルに保持。メモリ量を削減。



2. 直接数値シミュレーション（DNS）コードの概要

1) 非圧縮性一様等方乱流DNSコード（Box乱流）

領域 $\Omega = [-\pi, \pi] \times [-\pi, \pi] \times [-\pi, \pi]$ において

$$\frac{\partial u_i}{\partial x_i} = 0, \quad \frac{\partial u_i}{\partial t} + u_j \frac{\partial u_i}{\partial x_j} = -\frac{\partial p}{\partial x_i} + \frac{1}{\text{Re}} \frac{\partial^2 u_i}{\partial x_j^2} + f_i \quad \text{in } \Omega$$

u_i : 流速, p : 圧力, Re : レイノルズ数, f_i : 外力, 周期条件 on $\partial\Omega$

- 空間離散化：フーリエ・スペクトル法（3次元高速フーリエ変換）
aliasing errorは、 $\sqrt{2}N/3$ 以上のモードカットとphase shift法で除去
- 時間積分：4段4次ルンゲ・クッタ法

2) 等温を仮定した圧縮性乱流DNSコード

立方領域 Ω において

$$\frac{\partial \rho}{\partial t} + \frac{\partial \rho u_i}{\partial x_i} = 0, \quad \frac{\partial \rho u_i}{\partial t} + \frac{\partial (\rho u_i u_j + p \delta_{ij})}{\partial x_j} = \frac{\partial \tau_{ij}}{\partial x_j} + f_i,$$

$$p = \rho c^2, \quad \tau_{ij} = \mu \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} - \frac{2}{3} \frac{\partial u_k}{\partial x_k} \delta_{ij} \right) \quad \text{in } \Omega$$

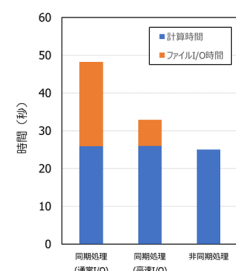
u_i : 流速, ρ : 密度, p : 圧力, c : 音速, μ : 粘性係数, f_i : 外力, 周期条件 on $\partial\Omega$

- 空間離散化
移流項：8次精度コンパクト差分
粘性項：8次精度中心差分
- 時間積分：3段3次TVDルンゲ・クッタ法
- 高波数領域における数値振動の除去：8次精度ローパスフィルタ

4. 予備実験 — チェックポイントファイルへの非同期入出力

時間発展熱伝導方程式を離散化したテストコードを作成し、途中経過を一定時間間隔でチェックポイントを取る場合について、SX-Aurora TSUBASAの非同期入出力機能による計算時間を計測した。

（計算時間）＞（I/O時間）の時、計算時間にファイルI/O時間が隠蔽されている。非同期入出力機構の正常動作を確認。



5. 今後の課題

- 非同期入出力機能の性能評価
 - ・ノード直結のSSDファイルシステム vs. 並列ファイルシステム
- 非圧縮性、圧縮性乱流DNSコードへの実装と評価

6. 参考文献

- Y. Takenaka et al., Optimizations of DNS Codes for Turbulence on SX-Aurora TSUBASA, Sustained Simulation Performance 2019 and 2020, pp. 51-59, Springer, (2021)
- Y. Sakurai et al., Effects of the Compressibility of Turbulence on the Dust Coagulation Process in Protoplanetary Disks, Astorophys. J., 911 (2), p. 140 (2021)
- T. Ishihara et al., Second-order velocity structure functions in direct numerical simulations of turbulence with RA up to 2250, Physical Review Fluids, Vol. 5, 104608, pp. 1-15 (2020)
- M. Yokokawa et al., I/O Performance of the SX-Aurora TSUBASA, Proc. of 2020 IEEE International Parallel and Distributed Processing Symposium Workshop (IPDPSW), pp. 27-35 (2020)

jh220043

Joint Usage / Research Center for Interdisciplinary Large-scale Information Infrastructures

平田 晃正 (名古屋工業大学)

熱中症リスク評価シミュレータの開発と応用

江川 隆輔 (東京電機大学) 滝沢 寛之 (東北大学サイバーサイエンスセンター)
堀江 祐圭, 木幡 咲英理 (日本気象協会), 小寺 紗千子, 高田 旭登, 木下 晃太郎, 植田 晴大 (名古屋工業大学)

JHPCN

研究背景と目的

- 熱中症による死亡者数は増加傾向。死亡例も報告。
- 本研究グループでは日本の夏場などの環境において、熱中症の主な要因である体温上昇、発汗量の解析を実施。
- 解析結果から熱中症のリスク評価を行い、熱中症予防の普及啓発活動に寄与することが目的。

解析手法

- 計算機上にて、様々な環境を模擬した仮想空間に人体モデルを配置、図1に示したフローチャートに従って体温変化や発汗量を計算。
- 温度上昇解析は生体熱輸送方程式を用いる。組織間の熱伝導や体表面から外気への熱伝達に加え、体温上昇に伴う発汗、血流量変化による熱輸送などの熱調整機能(システムバイオロジー)、太陽光吸収による熱発生を考慮し、時間領域有限差分法で逐次計算。

これまでの主な成果

- 2015年からの採択課題により高速化・並列化による準リアルタイム熱中症リスク評価システムを確立。ロードインバランスの解消等により、従来の2倍の解像度を持つ1mmモデルで3時間の暑熱ばく露における体温変化や発汗量を5.8分で試算可能。
- 名古屋市消防局と連携し、熱中症搬送者データを用いて高齢者の熱中症発症メカニズムについて考察。
- 長期間解析(2018-2020年6-9月)により得られた推定深部体温と実際に搬送された患者の搬送時体温とを比較、数日間の熱負荷蓄積が起因していることを示唆(図2参照)。

今年度の研究計画

- 熱中症搬送者数をはじめとする統計データと解析結果との連携
これまでに構築した熱中症搬送者数予測技術を、気象の異なる8都道府県に展開、その有効性を示す。解析期間は2013-2019年6-9月とし、入力変数には、一日の平均気温と、暑熱ばく露解析により算出した総発汗量、最大体温上昇値をそれぞれ用いる。
- 白内障発症メカニズム解明のための解析コードの改良と統計的考察
新たな応用展開として、これまで紫外線や外気温との関連性があるとされてきた核白内障発症のメカニズムについて、その発生要因を考察する。複合物理解析コードを改良、眼球(水晶体)の高精度推定を可能にし、核白内障有病率との相関を調査する。

解析結果・グラフ

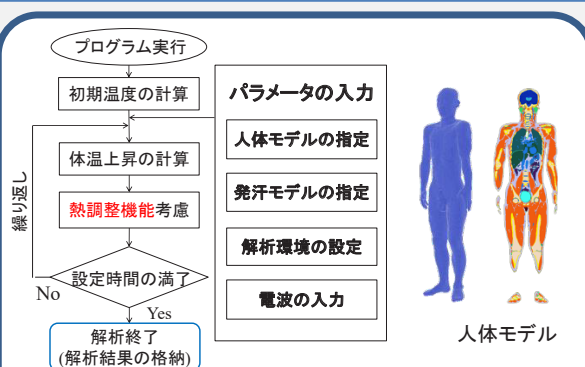


図1. 解析のフローチャート

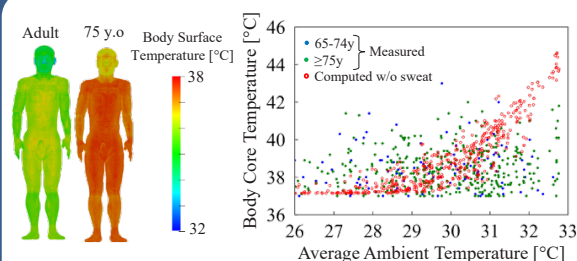


図2. 年代別温度分布の再現と熱中症搬送者(屋内・高齢者)の搬送時体温と解析値の比較

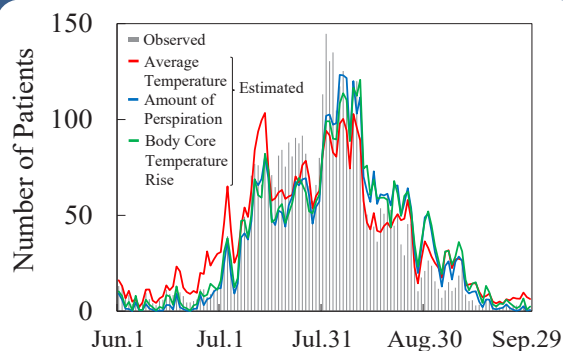


図3. 東京都における熱中症搬送者数予測(2013-2019年平均)

- 一例として、東京都における熱中症搬送者数の予測結果を示す(図3参照)。各入力変数において、高い予測精度を示しており(決定係数 $R^2 \geq 0.8$)、特に暑熱ばく露解析の結果を用いた予測で高い精度が得られた。

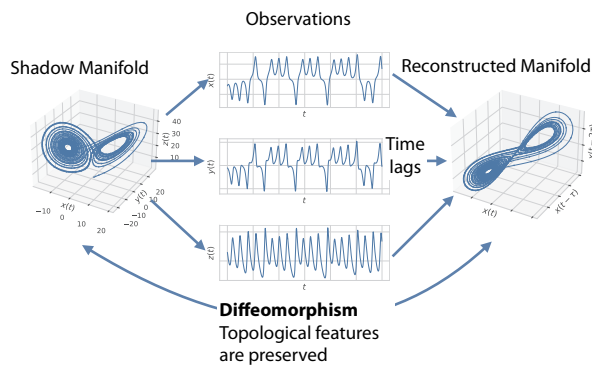
Project ID: jh220050 (国際課題)

Implementation and Application of High-Performance Empirical Dynamic Modeling

Keichi Takahashi¹, Gerald M. Pao², Wassapon Watanakesuntorn³¹ Cyberscience Center, Tohoku University, Japan keichi@tohoku.ac.jp² Salk Institute for Biological Studies, USA³ Nara Institute of Science and Technology, Japan

Background

We have been developing a high-performance implementation of Empirical Dynamic Modeling (EDM), **an emerging framework for non-linear time series analysis**, and applying it to large-scale datasets. EDM enables a variety of analyses such as short-term forecasts, quantification of non-linearity, and causal inference. These analyses are achieved by reconstructing the latent dynamics behind the data without assuming a parametric model or using prior knowledge (thus the name *empirical*).



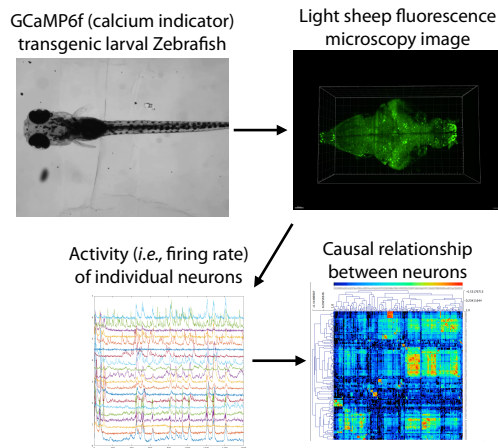
Objective

Although EDM is a generic method for modeling non-linear time series data, it was originally developed in the field of ecology. **Existing libraries for EDM analysis are thus designed to target a small number of short time series.** While recent studies have applied EDM to datasets including neural activity, gene expression and meteorology, the lack of a high-performance implementation limits the scale of datasets that can be analyzed. To handle large-scale datasets, we have been developing a high-performance EDM libraries [1, 2]. In this research, we continue this effort by tackling the following challenges:

- Porting our EDM implementation to the SX-Aurora TSUBASA vector engine:** Since EDM is a highly memory-bound algorithm, we will port EDM to the SX-Aurora TSUBASA vector engine to take advantage of its massive memory bandwidth. We plan to use our implementation of EDM named **kEDM** (<https://github.com/keichi/kEDM>) based on the Kokkos performance portability framework. 
- Enhancing the scalability of EDM by utilizing approximation algorithms:** A time-consuming kernel in EDM is the All k -Nearest Neighbor (AKNN) Search in the state space. Existing implementations use a brute-force algorithm, but this is clearly not scalable. We will thus integrate various approximate k -NN search algorithms (spatial trees, proximity graphs, product quantization, etc.) into EDM and evaluate which algorithm is best suited.
- Analyzing neural activity datasets to evaluate the performance of the ported implementation:** Our main goal is to perform EDM causal inference between neurons in an animal brain and elucidate how each individual neuron interacts with one another.

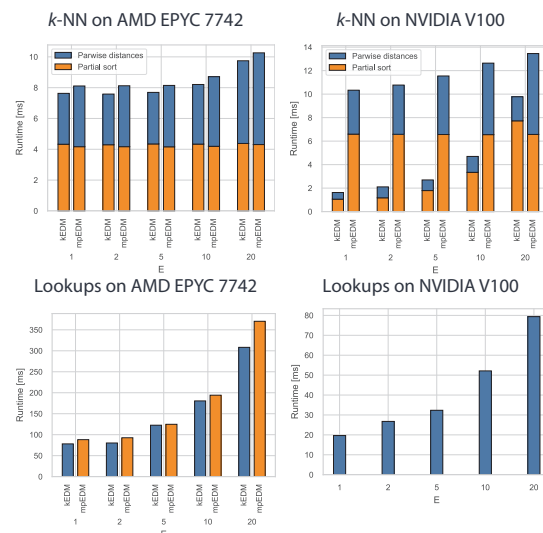
Target Application

We aim to analyze **the neural activity of an entire larval zebrafish brain that we recorded at single-neuron resolution** using light sheet fluorescence microscopy. So far, we identified a manifold that predicts turns of the fish at least 0.5 seconds ahead of time. Whenever the neural activity trajectory enters one of the loops of the manifold, the fish will turn.



Preliminary Performance Results

Using our EDM implementations **mpEDM** and **kEDM**, we have conducted preliminary performance measurements and roofline analysis on AMD EPYC 7742 and NVIDIA V100.



References

- [1] K. Takahashi, W. Watanakesuntorn, K. Ichikawa, J. Park, R. Takano, J. Haga, G. Sugihara, G. M. Pao, "kEDM: A Performance-portable Implementation of Empirical Dynamic Modeling using Kokkos", PEARC 2021, Jul. 2021.
- [2] W. Watanakesuntorn, K. Takahashi, K. Ichikawa, J. Park, G. Sugihara, R. Takano, J. Haga, G. M. Pao, "Massively Parallel Causal Inference of Whole Brain Dynamics at Single Neuron Resolution", ICPADS 2020, Dec. 2020.

salk
Where cures begin.

NAIST

