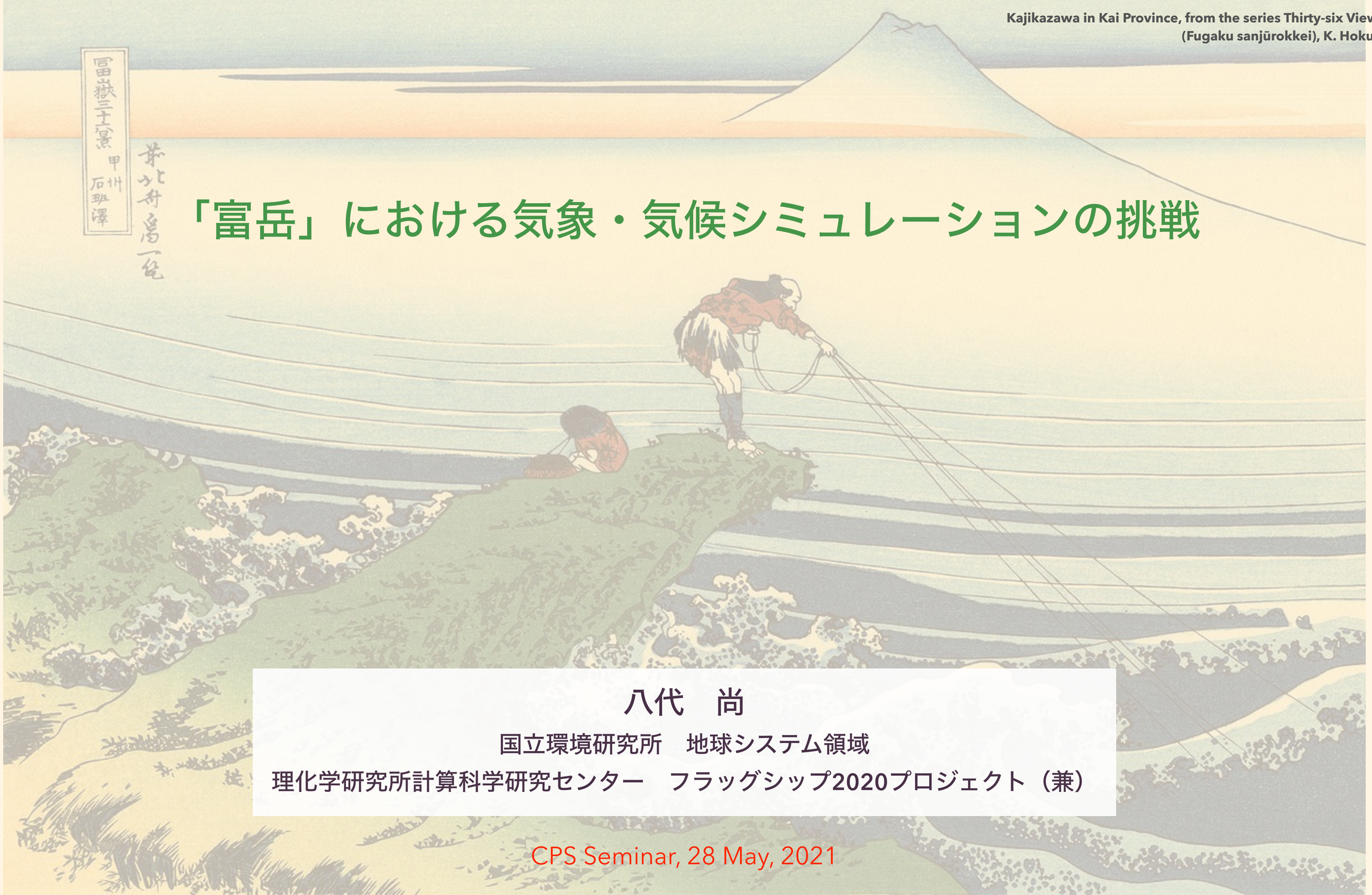


Kajikazawa in Kai Province, from the series Thirty-six Views of Mount Fuji
(Fugaku sanjūrokkei), K. Hokusai(1615-1868)

「富岳」における気象・気候シミュレーションの挑戦

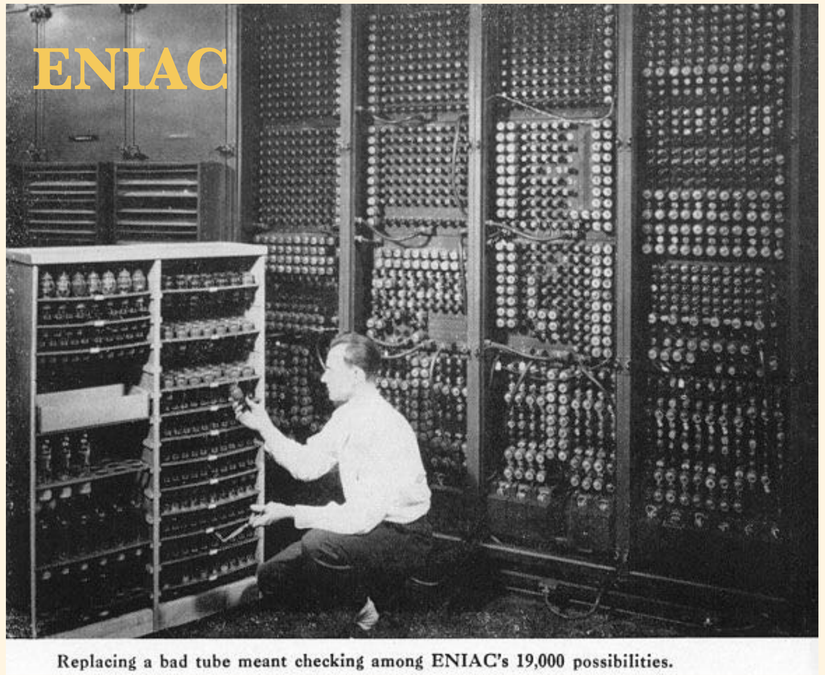
八代　尚

国立環境研究所　地球システム領域

理化学研究所計算科学研究センター　フラッグシップ2020プロジェクト（兼）

CPS Seminar, 28 May, 2021

気象・気候モデルとHPC



Replacing a bad tube meant checking among ENIAC's 19,000 possibilities.

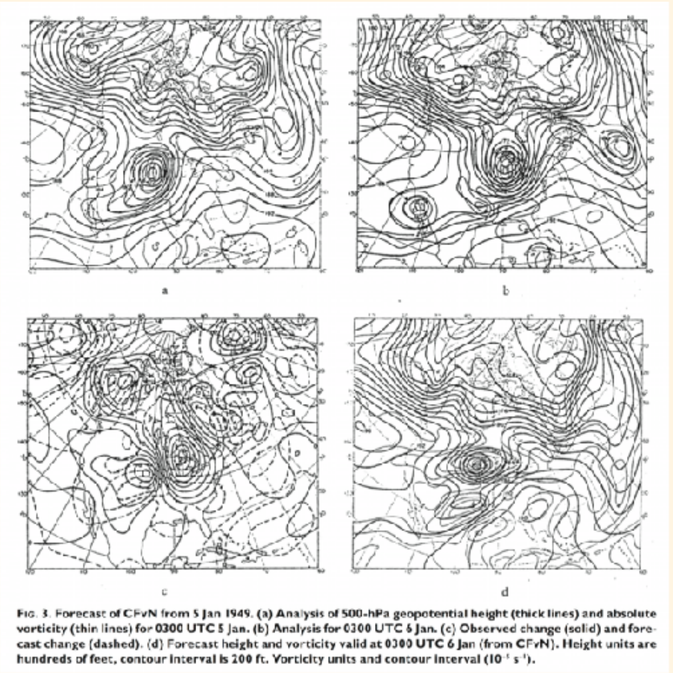
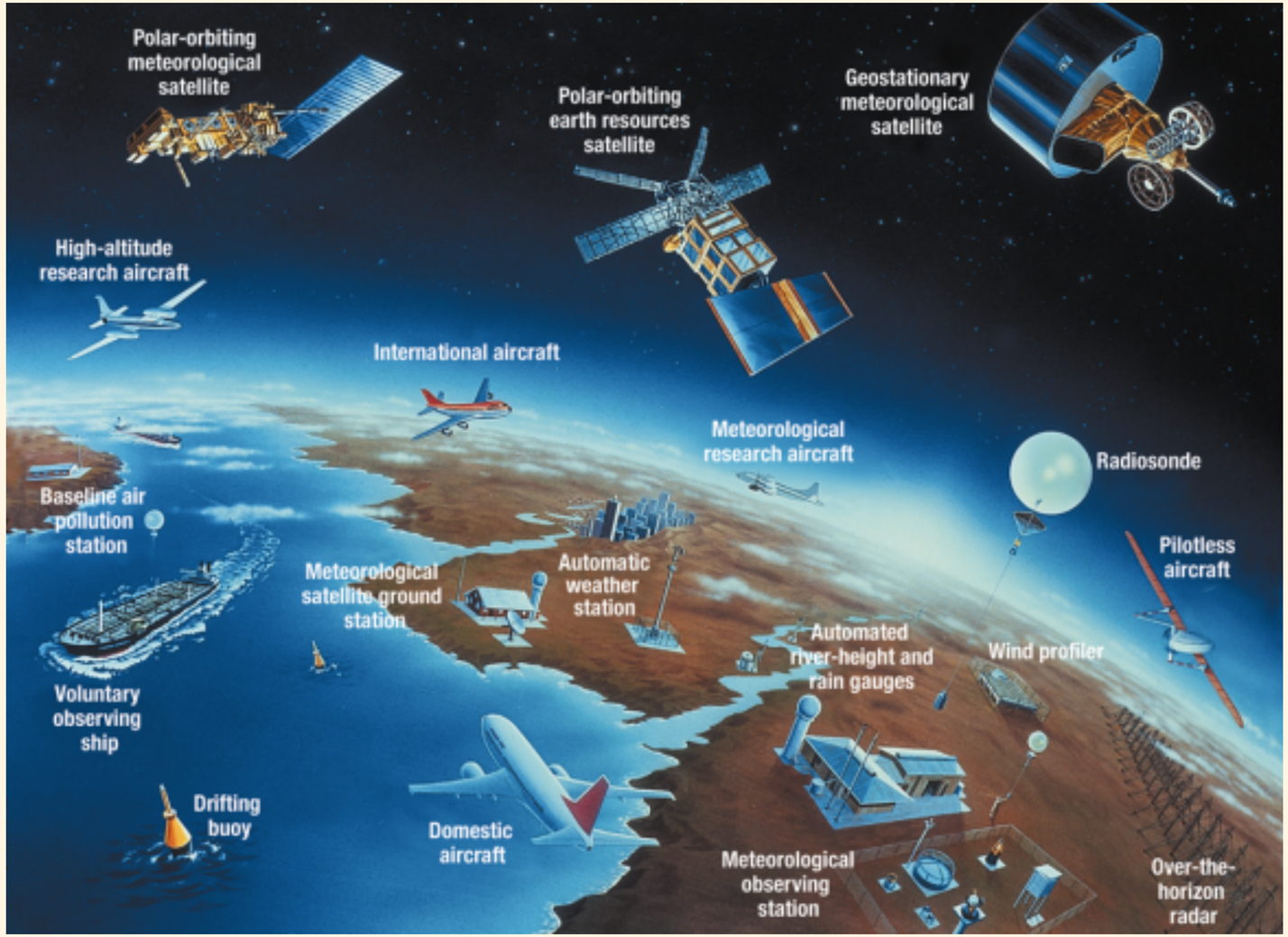


FIG. 3. Forecast of CFvN from 5 Jan 1949. (a) Analysis of 500-hPa geopotential height (thick lines) and absolute vorticity (thin lines) for 0300 UTC 5 Jan. (b) Analysis for 0300 UTC 6 Jan. (c) Observed change (solid) and forecast change (dashed). (d) Forecast height and vorticity valid at 0300 UTC 6 Jan (from CFvN). Height units are hundreds of feet, contour interval is 200 ft. Vorticity units and contour interval (10^{-5} s^{-1}).

- 気象学者はコンピュータユーザとしてはかなりの古株
  ：ENIACによる数値気象シミュレーション（1950）

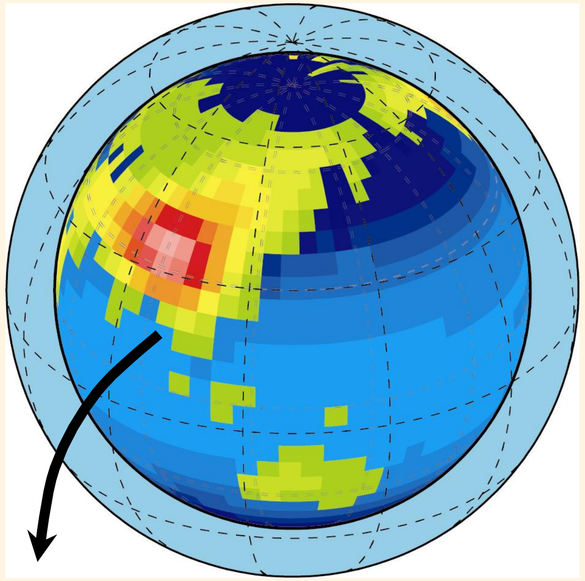
気象・気候モデルとビッグデータ



出典:WMO

- 気象学者はビッグデータのユーザとしてもかなりの古株
  ：時々刻々、様々なセンサによる観測結果が収集され、気象予報に役立てられている

計算科学的側面からみた気象・気候シミュレーションモデル（1）



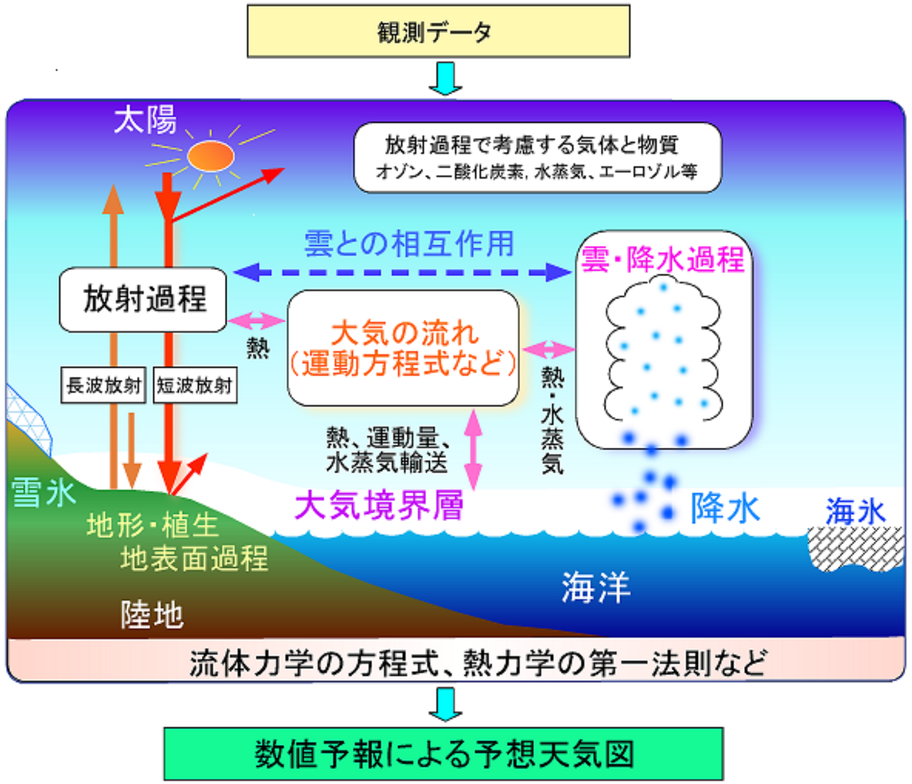
出典：気象庁

球殻状の3次元空間を離散化する

水平方向：緯度経度格子、準一様格子等
鉛直方向：高度座標あるいは気圧座標

スペクトル法：球面調和関数による展開
格子法：有限差分法、有限体積法

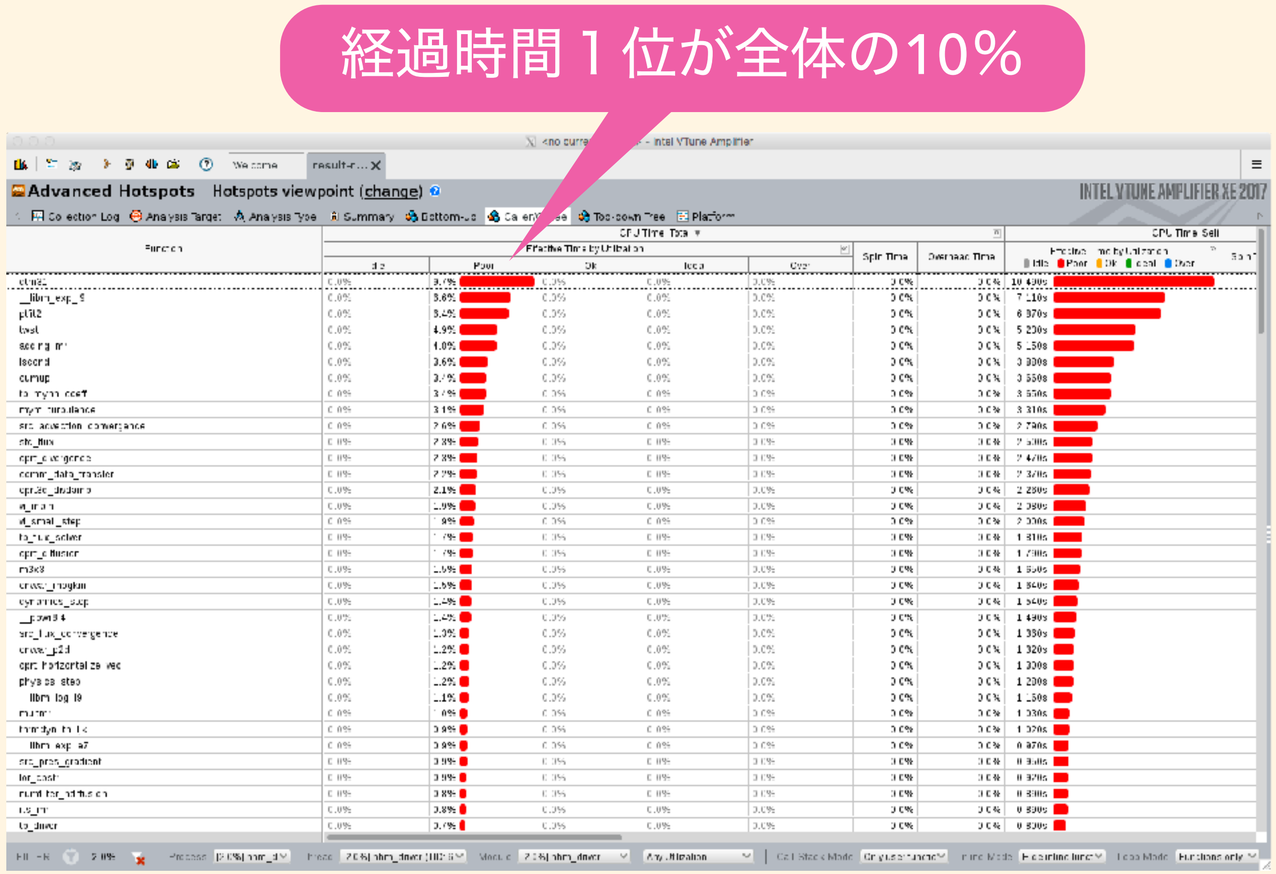
音波の扱い
：ブジネスク近似、非弾性近似、完全圧縮等
完全圧縮でも水平方向は陽解法、鉛直方向は陰解法で解くことが多い



出典：気象庁

- 気象計算は第一原理計算とは程遠い空間スケールで解いている：解くべき系の大きさが大きすぎるため
- 流体計算は一部に過ぎない：様々な物理的過程を解いている

計算科学的側面からみた気象・気候シミュレーションモデル（2）



Intel Vtune AmplifierによるNICAMのコストランキング結果

- “Flat profile”
  - 大部分の演算時間と演算量を占める区間(Hot spot)が存在しない
- Huge codebase
  - 数十万行にわたるソースコード
  - 学際的複合体：気象学の各専門分野から異なる演算パターンを持った数理モデルが持ち寄られている
- Huge memory footprint
  - 変数の数が非常に多い
  - メモリ律速、使い回しが効かない

Japanese flagship supercomputers

JAMSTEC, Yokohama

EARTH SIMULATOR

Earth Simulator (2002-2009)

RIKEN R-CCS, Kobe

京 K computer

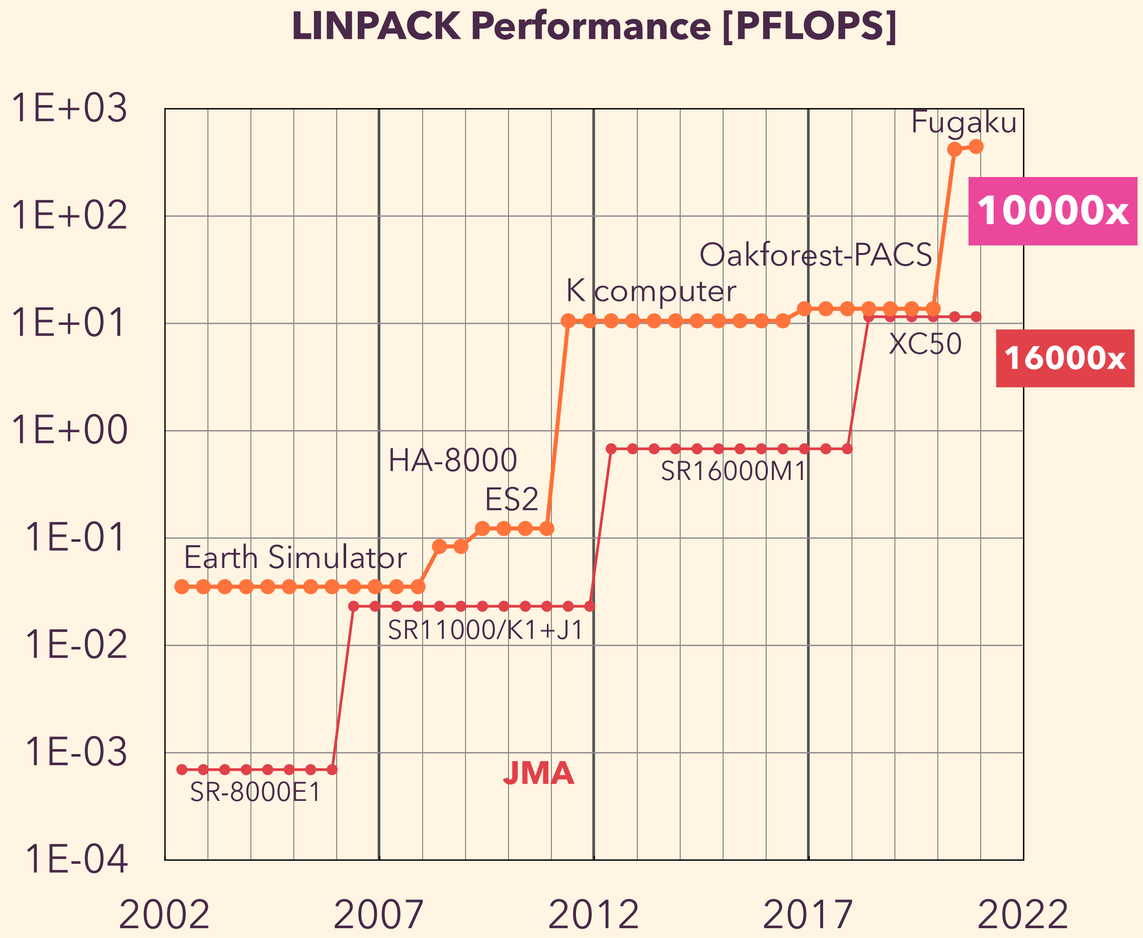
The K computer (2011-2019)

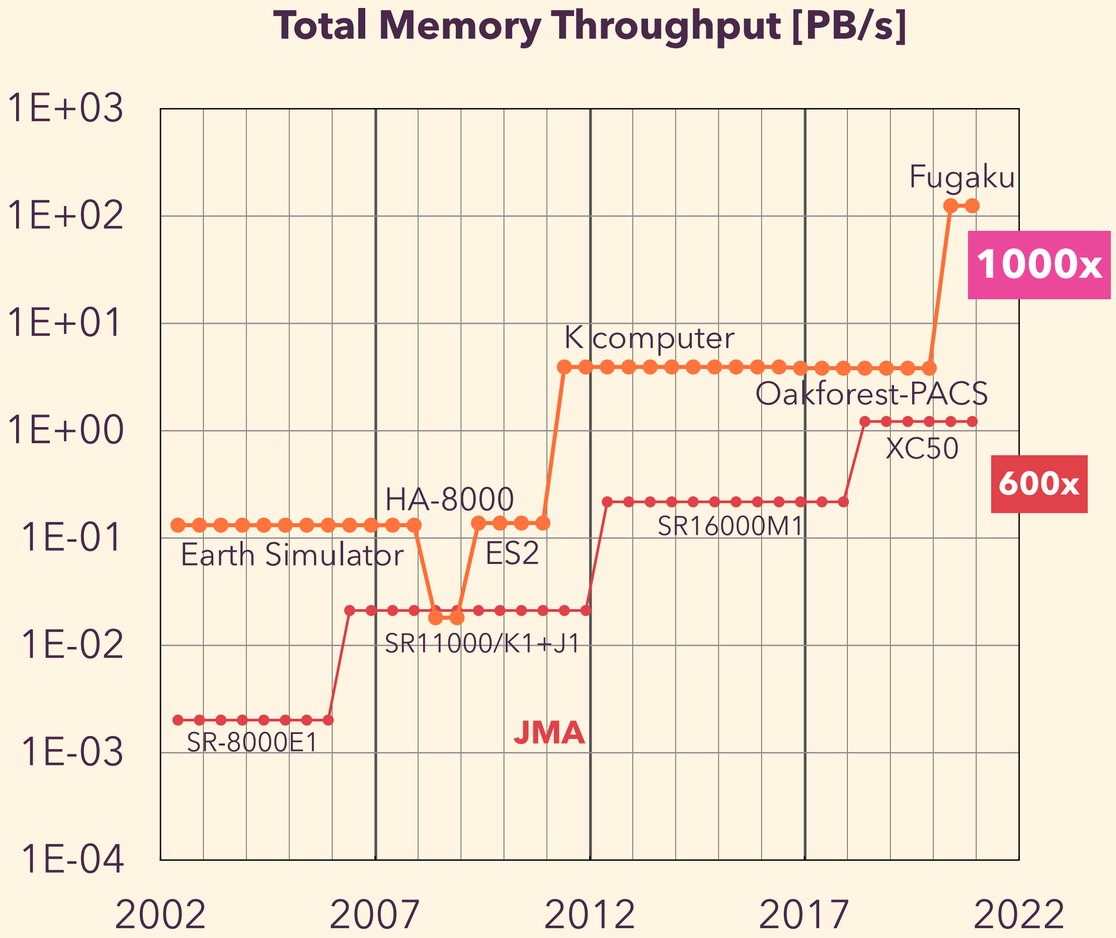
富岳

Fugaku (2020-)

RIKEN R-CCS, Kobe

日本の気象予報・気候研究が使ってきたスパコンの変遷

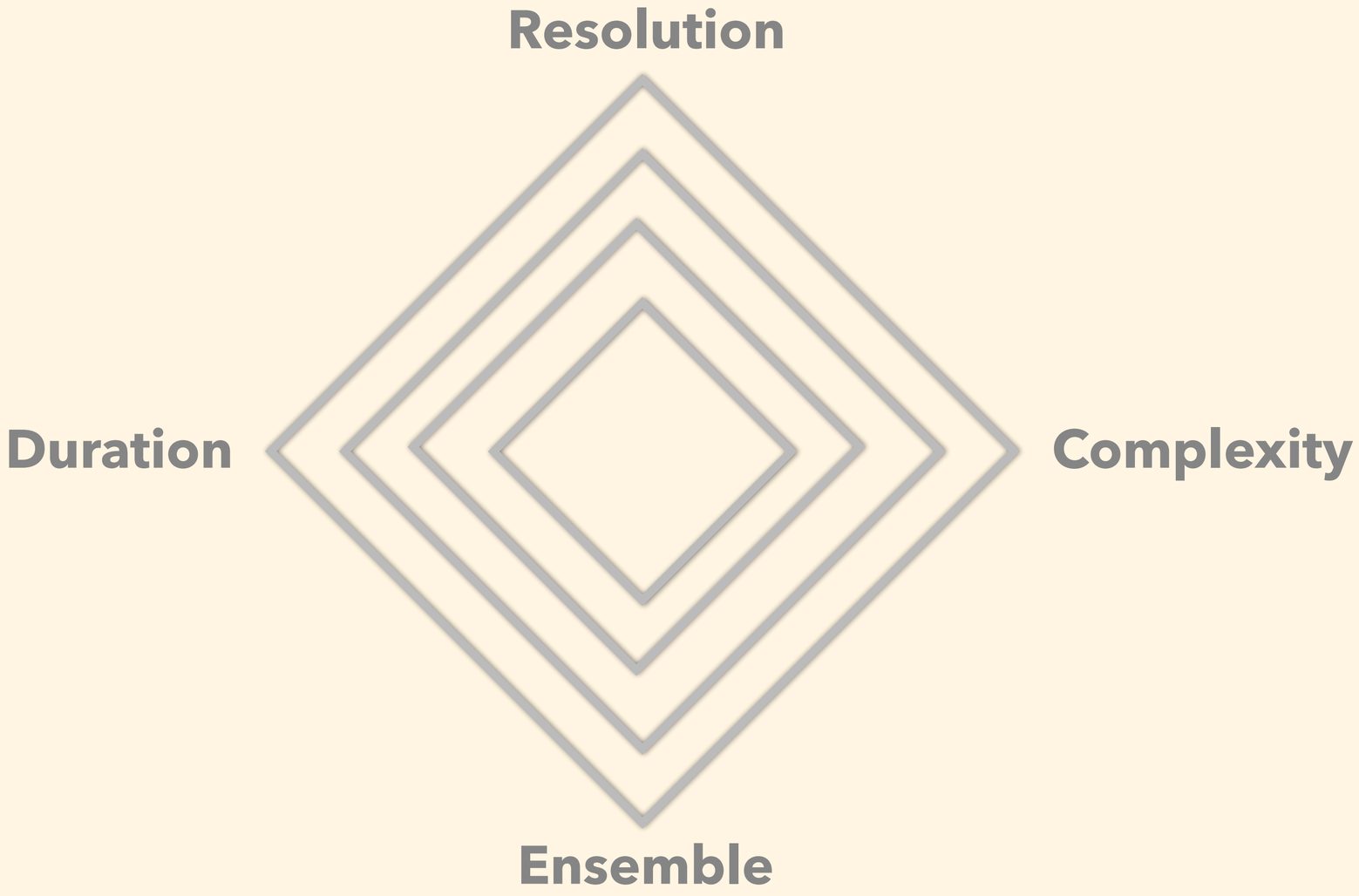




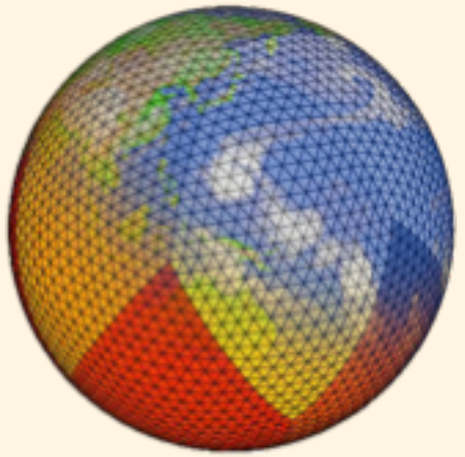
富岳と他のマシンの比較

	K	Fugaku	ECMWF 2020	ABCI	Earth Simulator (v4)
	SPARK64IIIVFX	A64FX	AMD EPYC 7742	NVIDIA Tesla V100	SX Aurora TSUBASA Type-20B
FLOPS	0.13 TFLOPS	3.4 TFLOPS	4.6 TFLOPS	7.8 TFLOPS	2.5 TFLOPS
Cores	8	48	128	5120	8
SIMD (for DP)	2	8	4	-	256
Mem. Cap.	16GB DDR3	32GB HBM2	256GB DDR4	16GB HBM2	48GB HBM2
Mem. throughput	0.06 TB/s	1 TB/s	0.4 TB/s	0.9 TB/s	1.5 TB/s
LLCache	6MB L2	8MB L2	256MB L3	8MB L2	16MB LLC
Byte/FLOP ratio	0.5	0.30	0.09	0.12	0.62
		FP16 support		FP16 support	PCI gen4

Demands on HPC resource for our research



正二十面体非静力大気モデルNICAM



- **2000年より開発開始**
  Tomita and Satoh (2005), Satoh et al. (2008, 2014)
  - 初代地球シミュレータを用いた世界初の3.5kmメッシュ全球シミュレーション
    Tomita et al. (2005), Miura et al. (2007, Science)
  - 京コンピュータを用いた世界初の0.87kmメッシュ全球シミュレーション
    Miyamoto et al. (2013, 2015), Kajikawa et al. (2016)
- **正20面体格子系・有限体積法・HEVI**
  - 球面調和関数展開のような大域通信を必要とせず、並列計算に適している
  - 準構造格子：各MPIプロセスが担当する計算領域は構造格子を保っており、間接参照による性能の低下が無い
  - 完全圧縮流体、静力学近似を用いない、質量とエネルギーの高精度な保存

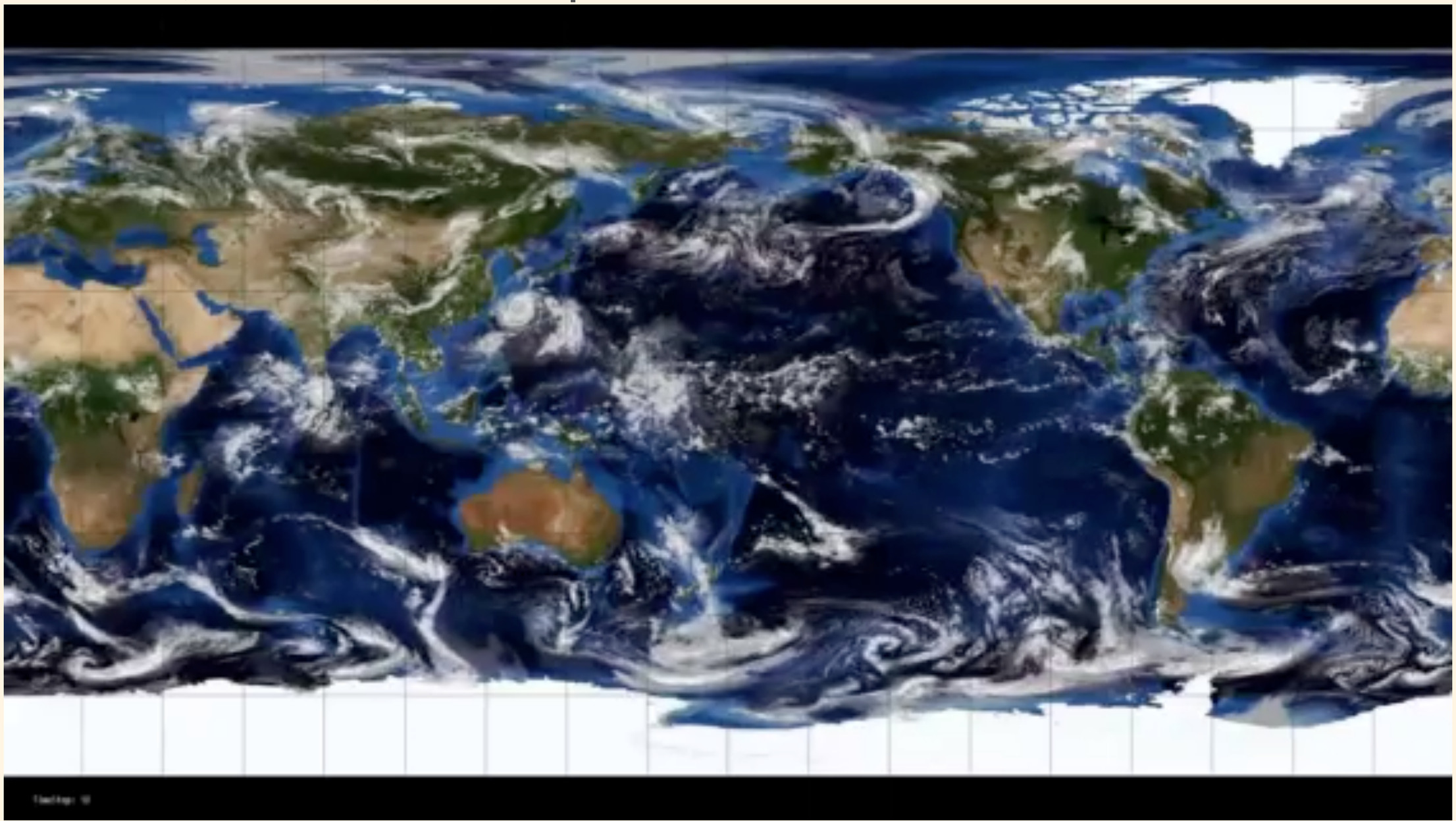
Prof. Masaki Satoh
(AORI, Univ. of Tokyo)

Dr. Hirofumi Tomita
(RIKEN R-CCS)

The first global sub-km weather simulation (Miyamoto et al., 2013)

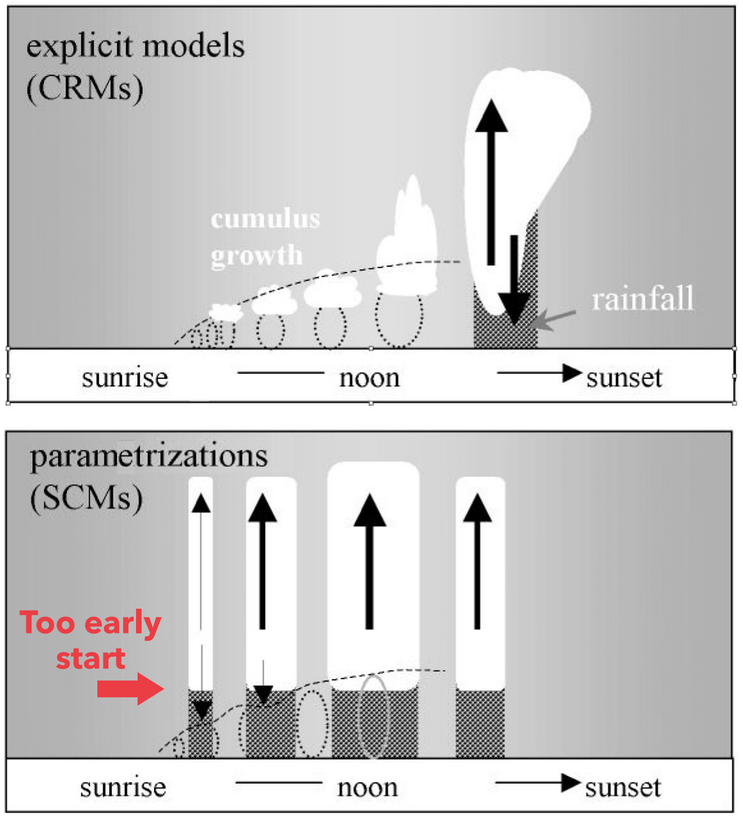
20480nodes(163840cores) on the K computer

Visualized by Ryuji Yoshida(NOAA/CIRES)



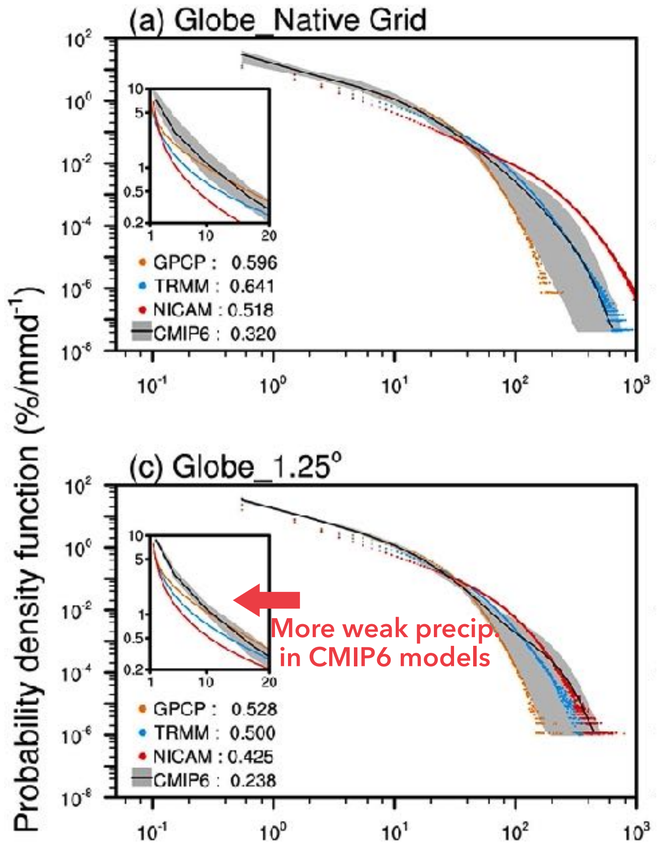
Cloud-permitting simulation: the big leap

Diurnal cycle of deep convection



Guichard et al. (2004)

Global precipitation PDF

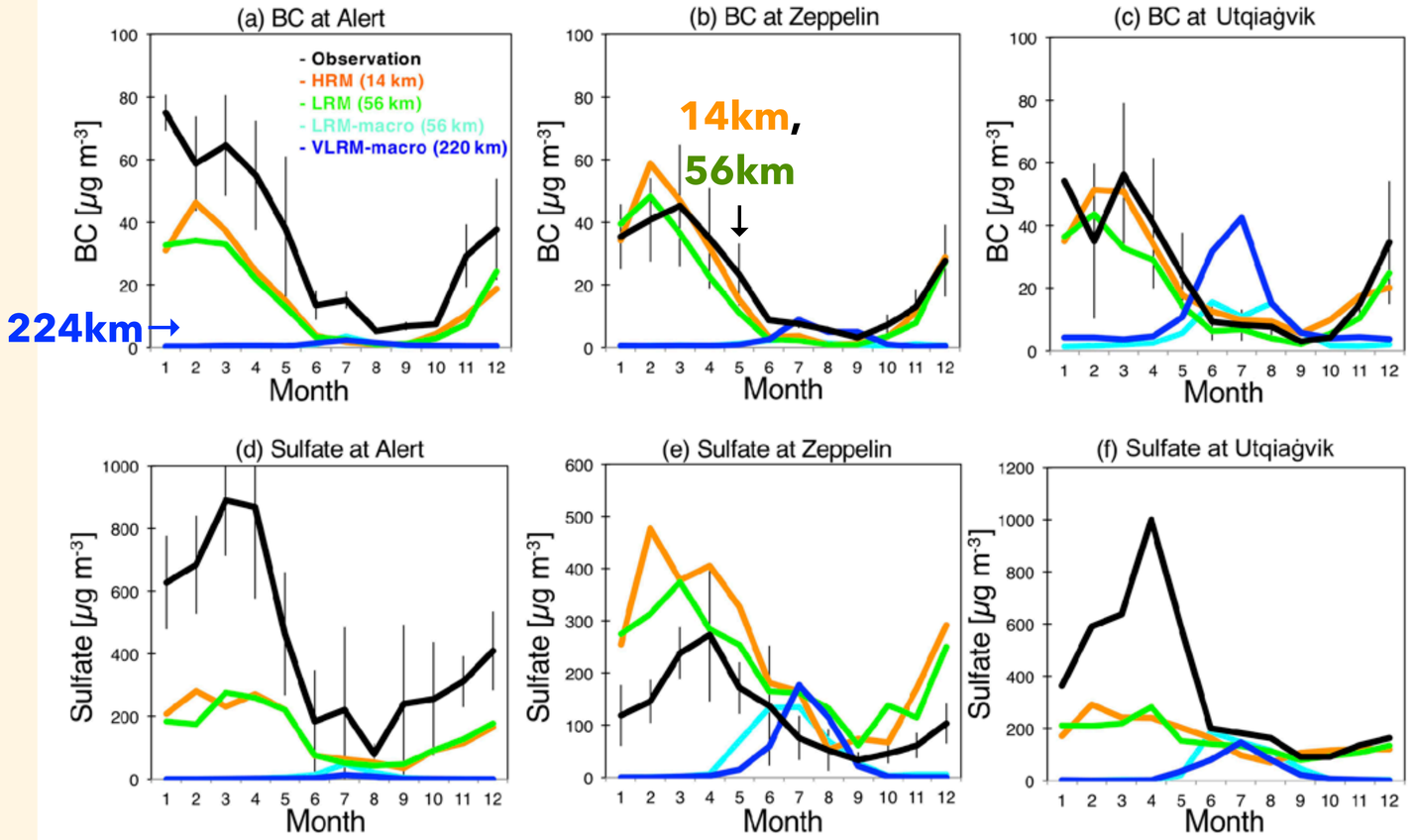


Na et al. (2020)

Cloud-permitting simulation: the big leap

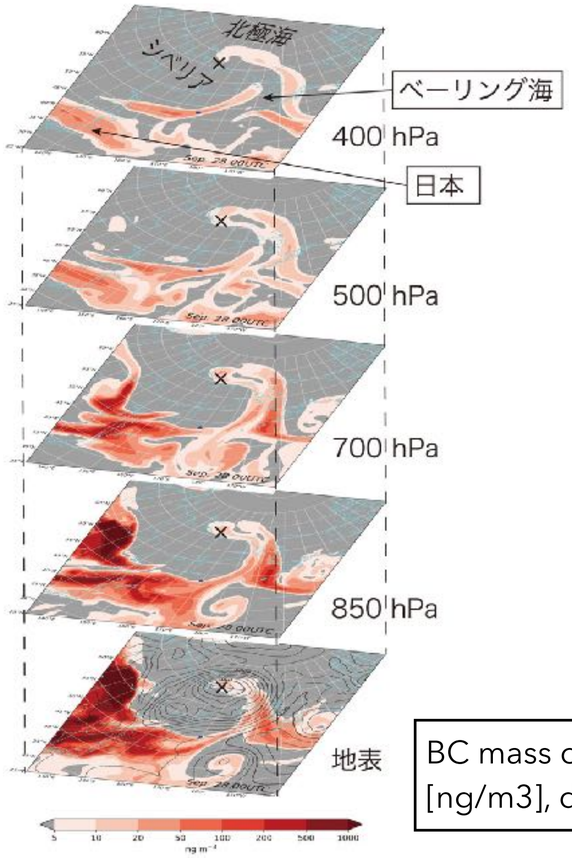
for environmental studies

Seasonal cycle of aerosol concentrations



Goto et al. (2020)

BC intrusion to the Arctic



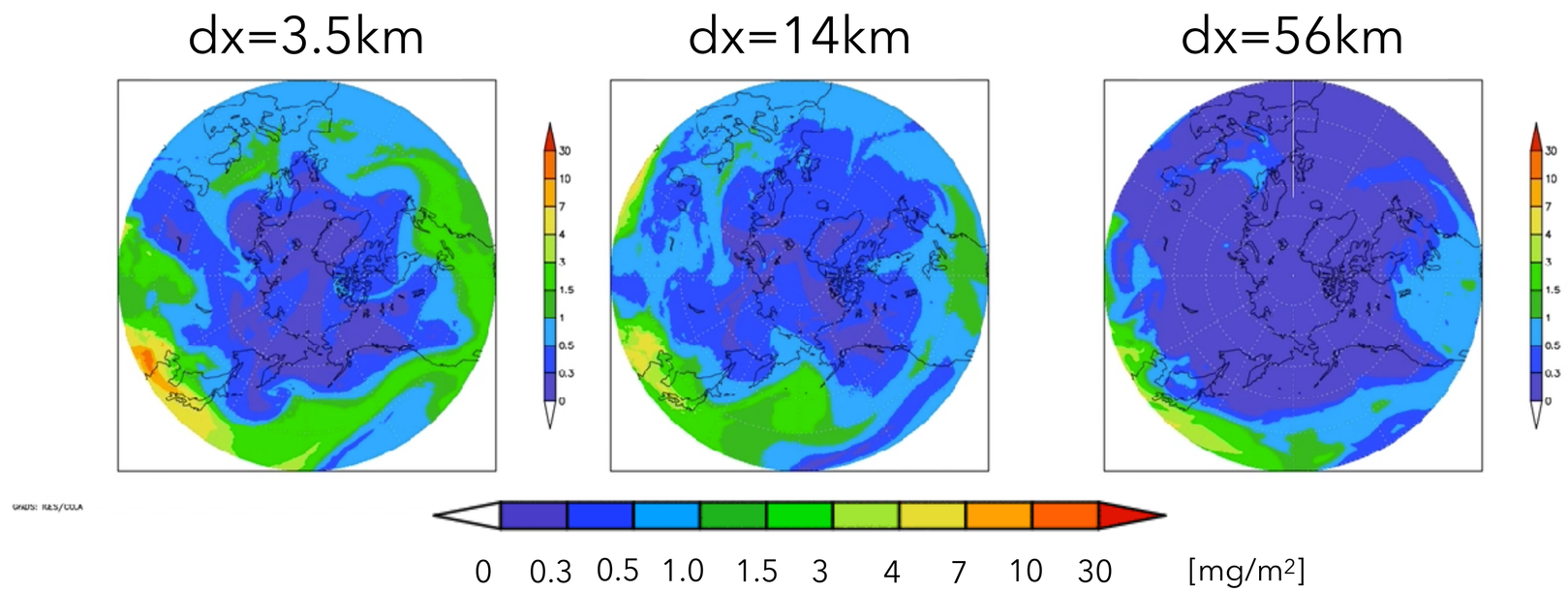
BC mass concentrations [ng/m3], dx=56km

Yamashita et al. (2021)

Cloud-permitting simulation: the big leap

for environmental studies

BC intrusion to the Arctic

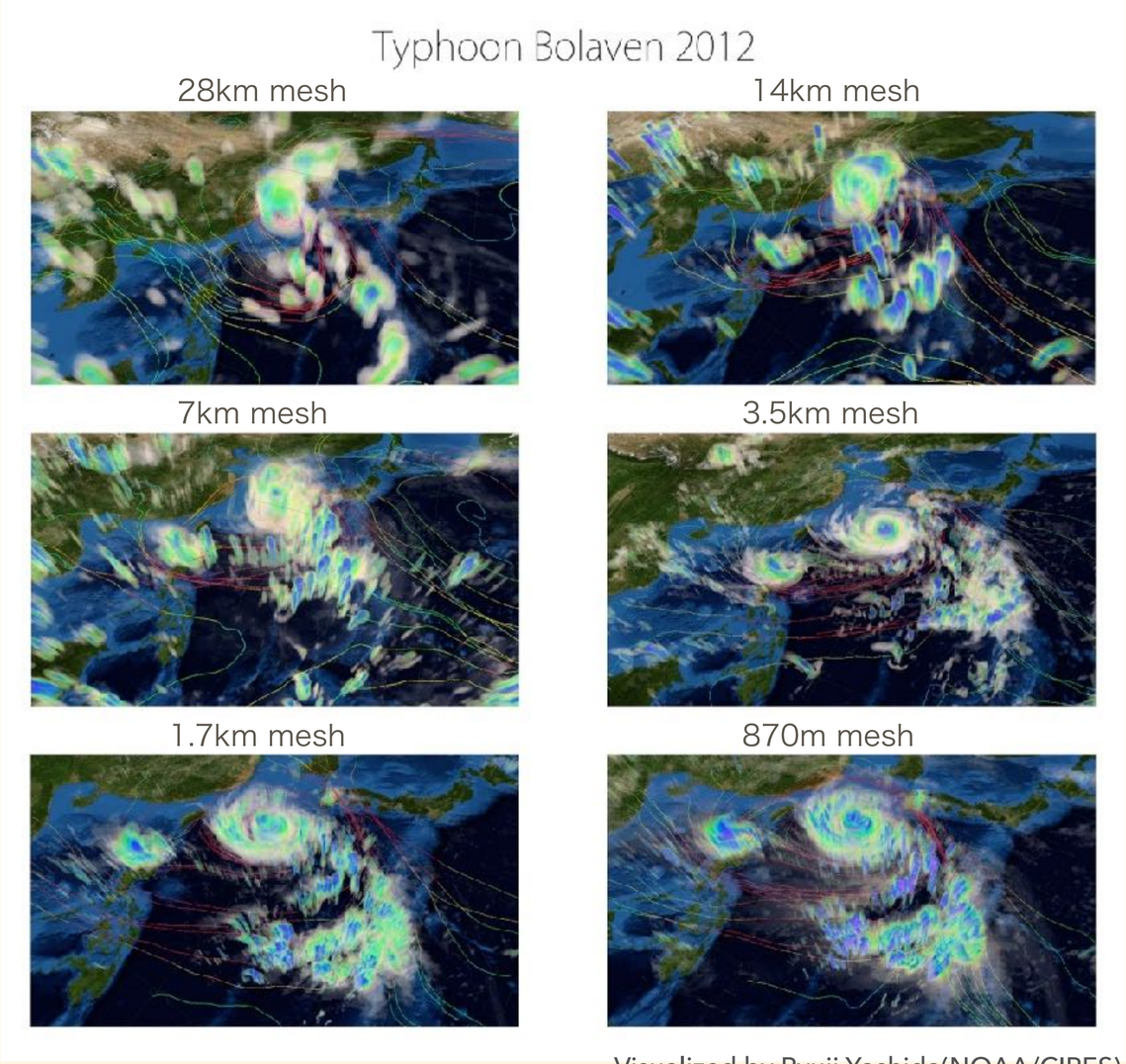


Vertically accumulated BC in the Arctic region (10 days)

Sato et al. (2016), Scientific Reports

- Amount of BC transported to Arctic is increased with fining resolution
- Mid-latitude low and frontal system has large contribution to the BC transport

High-resolution simulation is expensive!



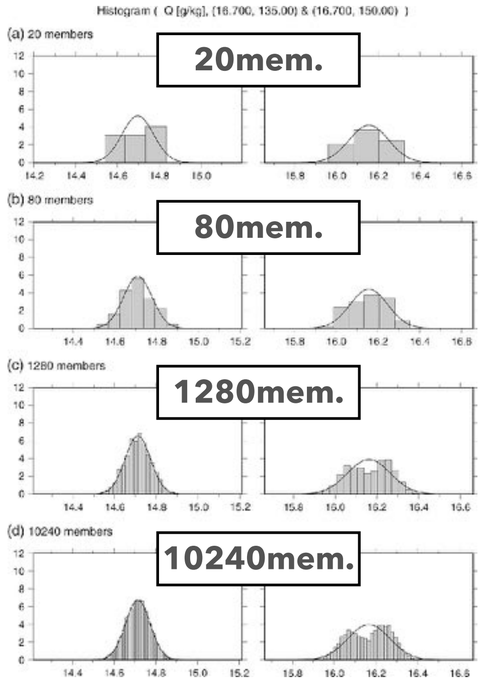
Visualized by Ryuji Yoshida(NOAA/CIRES)

- 格子点数とシミュレーションの再現性
  - 格子の数が多い（解像度が高い）ほど、より細かいスケールの現象を捉えることができる
  - 簡略化していた過程をちゃんと表現できるようになる
  - ただし、計算量は爆発的に増える：水平解像度を倍にすると計算量は８倍

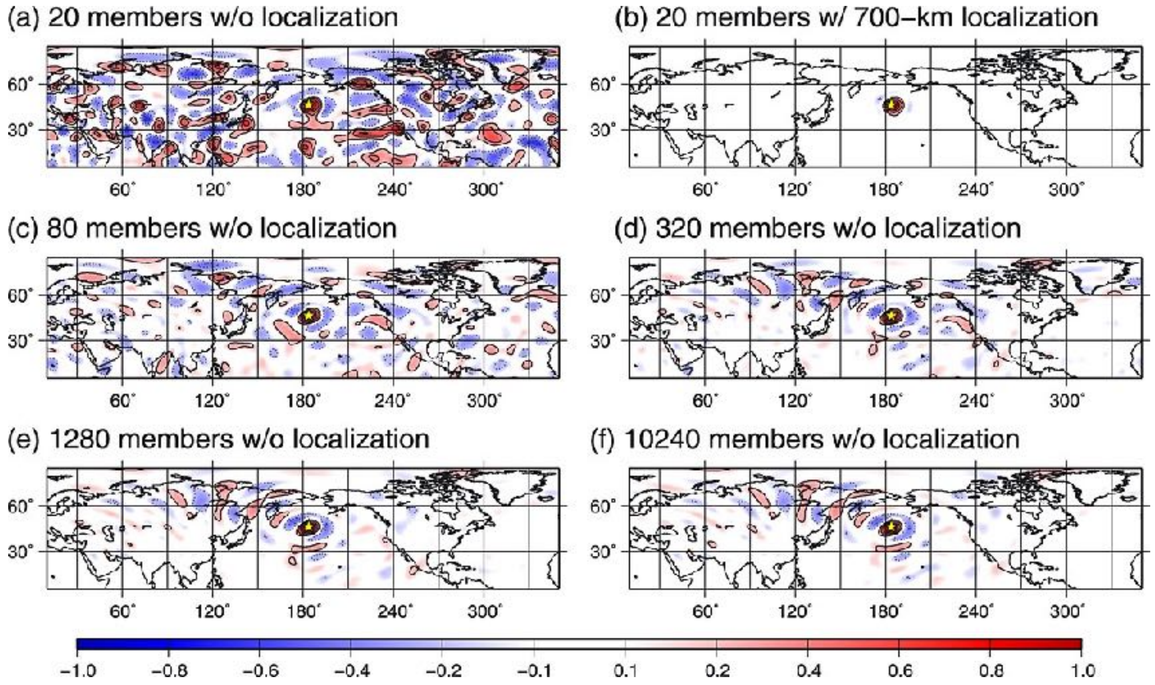
x30,000 calculation cost!

The impact of large ensemble in data assimilation

Histograms for water vapor



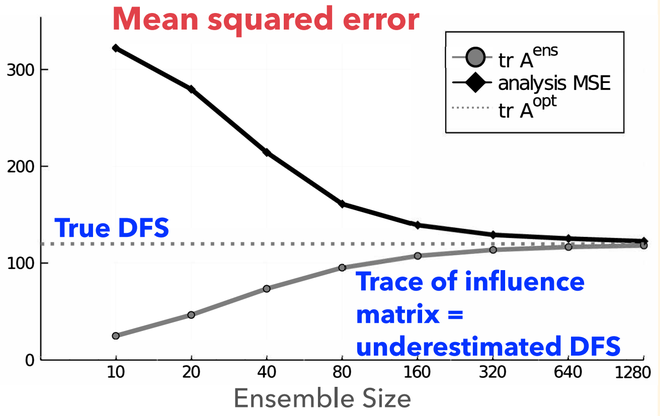
Ensemble-based autocorrelations



Miyoshi et al. (2014)

Non-localization case

DFS = Degrees of Freedom for Signal



Hotta and Ota (2021)

富岳プロジェクトにおける我々の取り組み

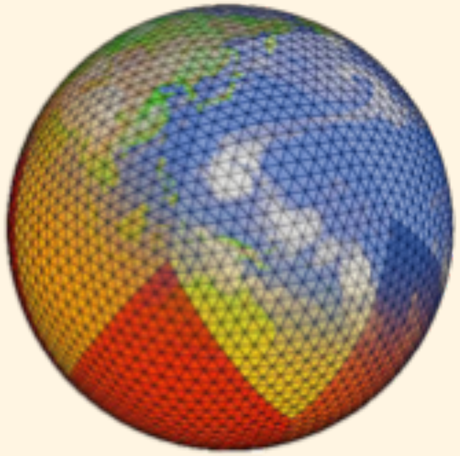
気象・気候分野のシステム-アプリケーションコデザイン

ポスト「京」重点課題４：
観測ビッグデータを活用した気象と地球環境の予測の高度化

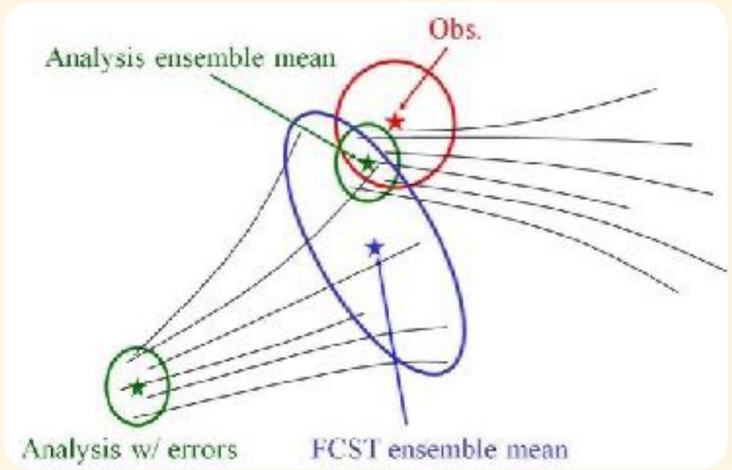
理研R-CCS
FLAGSHIP2020プロジェクト

Target application in co-design process:
NICAM+LETKF

NICAM+LETKF



+



NICAM（全球高解像度大気モデル）

LETKF（アンサンブル同化システム）

京での成果：世界最高解像度

(水平870m)での全球大気シミュレーションに成功

どのような特性を代表しているか

- 構造格子ステンシル計算
  ：メモリ間接参照は無いが、メモリバンド幅で律速される
- 物理諸過程を解くコンポーネントは、ループ内の行数が非常に多く複雑なコード
  ：コンパイラによる最適化が難しい

京での成果：世界最大規模

(10240メンバー)での全球大気データ同化に成功

どのような特性を代表しているか

- ストレージ・通信を用いた膨大な量のデータ交換
  ：ファイルI/O性能、大域通信性能を要求
- １プロセス内に収まる小さいサイズの固有値計算を大量に反復する

エクサスケールはトレードオフの時代

- 数理モデル発展期（1980年代〜1990年代）：貧弱な計算性能でも精度高い定式化
- 演算性能高度成長期（2000年代〜2010年代）：計算リソースを如何に使い切るか
- 現代（2020年代〜2030年代）：計算機の多様化、多層化、複雑化

- 「安くて」「速くて」「ソフトウェア開発コストの低い」スパコンは無い
  - 何らかのトレードオフを意識して、最適解を探さねばならない：その過程が**コデザイン**
  - エクサスケールの時代はよりその傾向が強い：省エネ性、開発費用の増大、等
  - 富岳のコデザインは世界でも特筆すべきコデザインの成功例であると私は考える
    - アプリケーションがCPU開発段階から物申すことが出来た
    - コンパイラ開発者、数値ライブラリ開発者との協働

スーパーコンピュータ「富岳」とコデザイン

- アプリケーション性能で「京」の100倍を目指す
  - ならば、プロダクションに直結する（＝気象予報で運用するに耐えうる）複雑度でのアプリケーション評価をしようじゃないか
  - いろいろ理想的に簡略化したアルゴリズムで速くても実用的な意味はない
  - 気象予報は予測シミュレーションだけではない
  - 最適な初期値を観測データと組み合わせて作る「データ同化」もベンチマークの対象に
- どうやって100倍を達成する？
  - 演算性能は1CMGあたり6.6倍、メモリ性能は4倍、利用可能プロセス数はおよそ6.4倍→足りてない
  - CPUからストレージまで、システム全体を使ったワークフロー全体の高速化を目指す
  - さらに、シミュレーションを高速化する大きな方針転換が必要：単精度浮動小数点演算の積極的利用

アプリケーション側のコデザイン努力

NICAM

Source code refactoring

Loopwise optimization

w/ compiler development

single precision

Nakano et al. (MWR,2018)

Speedup
Pre-/Post-process

Data exchange method

Yashiro et al. (GMD,2016)

LETKF

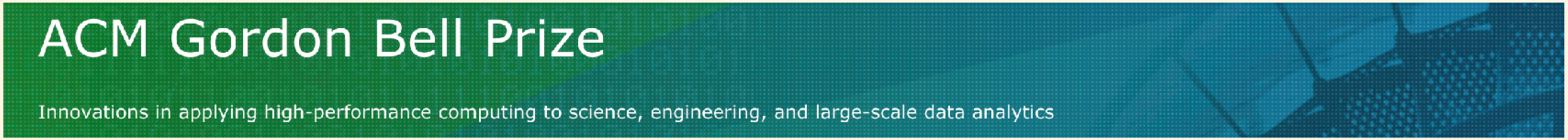
Big observation data

Eigenvalue solver

w/ numerical library development

Source code refactoring

グランドチャレンジラン

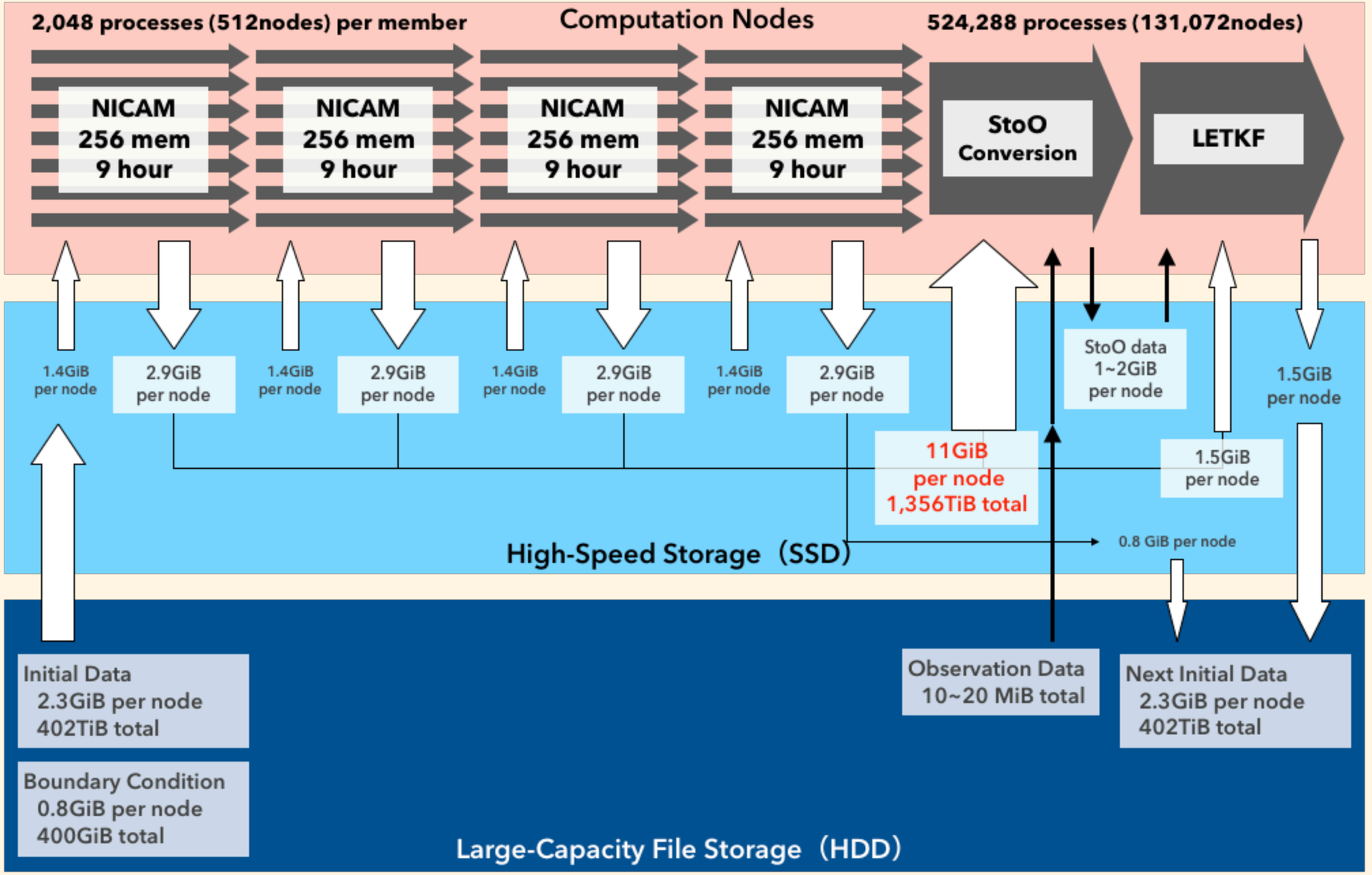


2020 Finalist paper

H. Yashiro, K. Terasaki, Y. Kawai, S. Kudo, T. Miyoshi, T. Imamura, K. Minami, H. Inoue, T. Nishiki, T. Saji, M. Satoh, and H. Tomita, "**A 1024-Member Ensemble Data Assimilation with 3.5-Km Mesh Global Weather Simulations**," *in SC20: International Conference for High Performance Computing, Networking, Storage and Analysis (SC)*, Atlanta, GA, US, 2020 pp. 1-10.

富岳の開発目標である「アプリケーション性能で100倍」を、NICAM-LETKFを用いて、
ターゲット問題である全球3.5km、1024メンバーアンサンブルデータ同化で達成

NICAM＋LETKFデータ同化実験でのデータフロー



富岳での性能最適化の中で現れたトレードオフ（1）

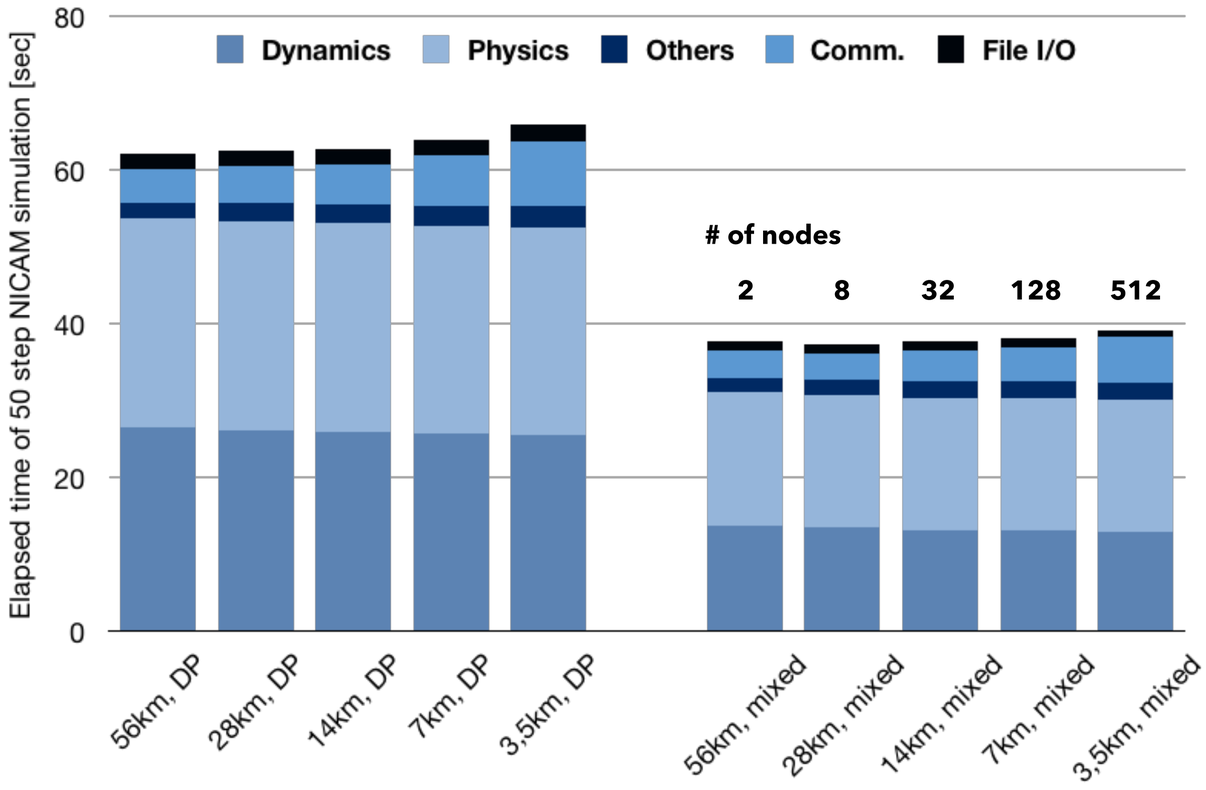
- ループレベルの最適化
  - 並列性の確保 vs データ局所性の確保
    - 京では2SIMD、8スレッド並列→富岳では8SIMD(倍精度)、16SIMD(単精度)、12スレッド並列
    - AoSoA型のループ構造が必要：データ構造は地球シミュレータ時代より引き継ぐSoAで
    - 最終的にはメモリ転送性能がものを言う：富岳の高いB/F比（~0.3）が全体を支えている
  - データ局所性の確保 vs レジスタスピル
    - 富岳ではレジスタサイズが大幅に減少、データ局所性向上のための中間配列の削減が更に追い打ちをかけた
    - 再度ループ分割を行えばよいが、最適な分割数はマシンパラメータに拠る
    - コンパイラ開発との協働による自動ループ分割の適用

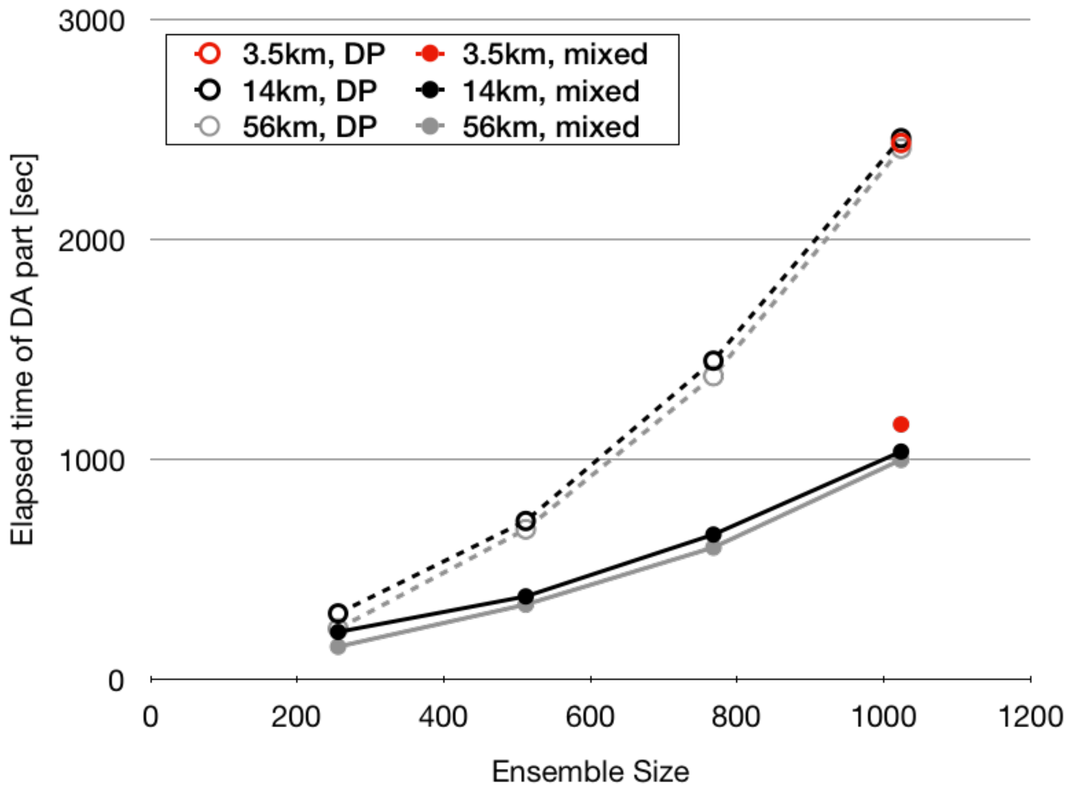
富岳での性能最適化の中で現れたトレードオフ（2）

- 単精度浮動小数点計算の積極的な利用
  - 演算性能 vs シミュレーション性能
    - 力学コアの理想実験における検証（Nakano et al., 2018）：モデル間の差よりも誤差は小さく抑えられる
    - サブルーチン単位での精度担保と自動チューニングへ発展：名古屋大片桐先生との共同研究
  - 演算性能 vs データ同化解析性能
    - LETKFの計算コアであるGEMMと固有値計算の単精度化：理研今村T・工藤さんによるKevdライブラリ

- NICAMとLETKF間の最適化
  - 演算量のインバランス vs データ移動量
    - NICAMの空間分割手法をLETKFにも適用し、ファイルI/Oのノード内局所化、ローカルファイルシステム（SSD）の利用、大域通信のグループ化を行うことにより、劇的なLETKFの高速化を実現：Throughput-awareなアプリケーション間連携
    - 各PEの演算量は不均一に分布する観測データをいくつ同化するかによって変わる
      ：動的なロードバランシングよりもデータを「なるべく動かさない」ことを選択した

Performance results on Fugaku





NICAM part

- Weak scaling: same problem size per node
- x1.6 speed-up by mixed (DP-SP) precision

LETKF part

- Data-centric design reduced file I/O time drastically
- Optimized Eigenvalue solver w/ single precision saved computation time in large ensemble case

Cycle test of NICAM+LETKF

A. 14-km Mesh, 1024-member

	Time [sec]	Computation [TFLOPS]	Memory Throughput [TB/s]
One DA Cycle total	**5,007**	**2,589**	**2,858**
Simulation part total	**3,973**	**1,824**	**3,517**
NICAM set1	1,026	1,766	3,404
NICAM set2	982	1,845	3,556
NICAM set3	981	1,847	3,561
NICAM set4	984	1,841	3,550
DA part total	**1,034**	**5,527**	**326**
StoO	83	77	99
LETKF	951	6,003	345

B. 3.5-km Mesh, 1024-member

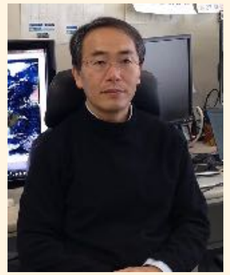
	Time [sec]	Computation [TFLOPS]	Memory Throughput [TB/s]
One DA Cycle total	**14,200**	**33,234**	**54,990**
Simulation part total (estimated from set1 x 4)	**13,042**	**29,174**	**59,459**
NICAM set1 (estimated from shorter time steps)	3,260	29,174	59,459
DA part total	**1,158**	**78,970**	**4,653**
StoO	196	522	671
LETKF	961	94,995	5,466

- Fully coupled experiment for 14-km mesh with 8,192nodes
  - Less than 1.5 hours for one DA cycle
- Quasi-coupled for 3.5-km mesh with 131,072 nodes (82% of the full nodes)
  - ~4 hours for one DA cycle: **100x** faster than the estimated time on the K computer

「富岳」成果創出加速プログラム

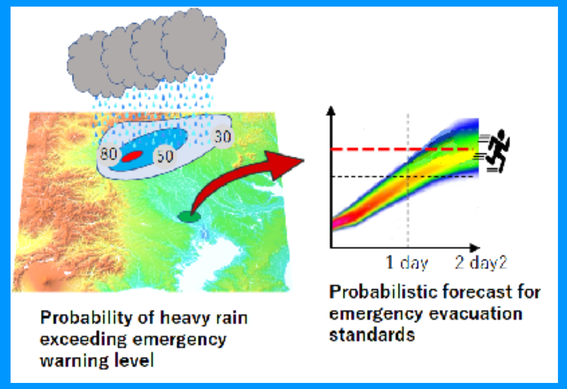
「防災・減災に資する新時代の大アンサンブル気象・大気環境予測」

(PI:佐藤正樹教授, FY.2020-2022)



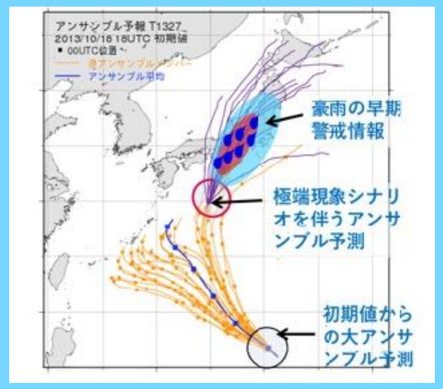
(co-P.I. Takemasa MIYOSHI, RIKEN)

キーワード: 大アンサンブル



NRI-NHM
ASUCA
SCALE-RM
Debris flow model

Theme1: Short-range regional prediction
(P.I. Takuya KAWABATA, MRI)



NICAM
NICAM-COCO
Typhoon, MJO

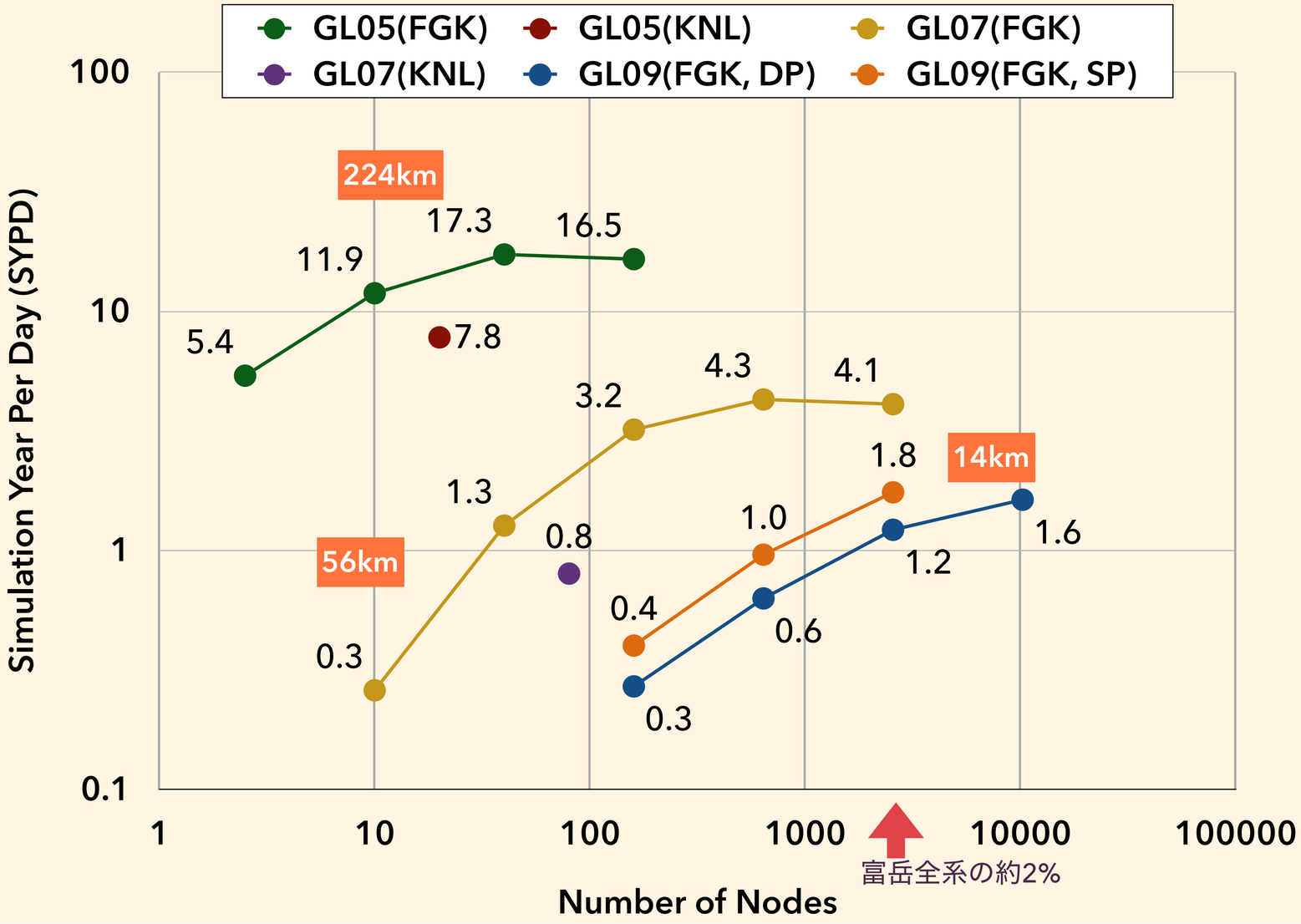
Theme2: Global-scale prediction
(P.I. Tomoki MIYAKAWA, AORI)

NICAM-Chem
NICAM-TM
GHG, SLCF, SWI

Goto, Uchida, Niwa, Yamashita, Tanoue

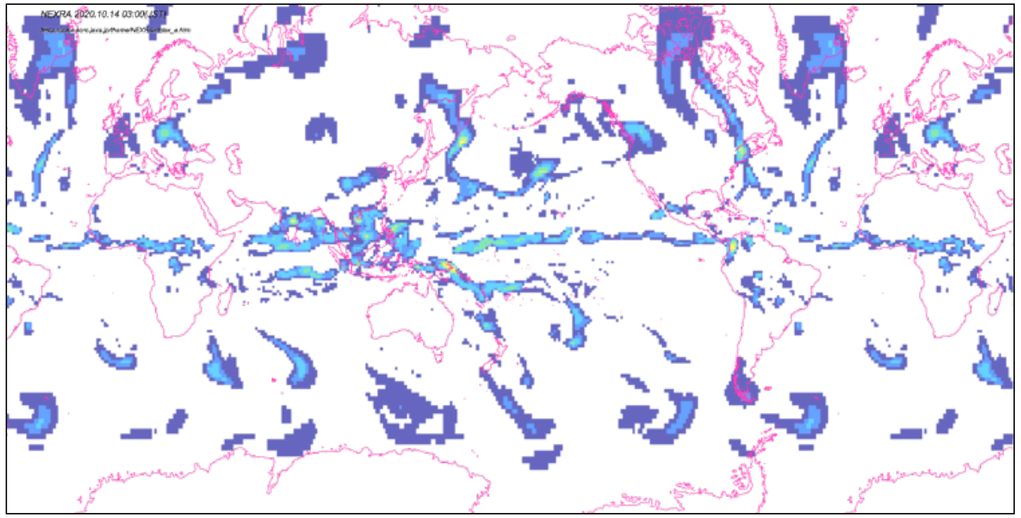
Theme3: Advanced technology of data assimilation
(P.I. Hisashi YASHIRO, NIES)

富岳ではNICAMはどのくらい計算できる？



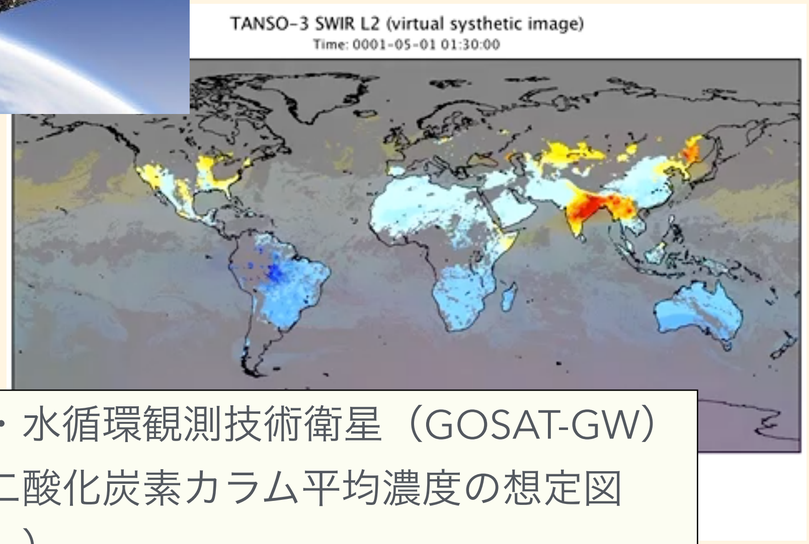
- NICAM.19.1で計測
- すべて78層
- GL05~07: LSC, Chikira
- Δt [sec]
  - GL05: 720, MP720, RD3600, SFC120, TB60
  - GL07: 120sec, MP120, RD1800, SFC60, TB30
  - GL09: 30sec, MP10, RD1800, SFC30, TB30

成果の社会実装（テーマ３担当部分）



NICAM-LETKF JAXA Research Analysis (NEXRA)

温室効果ガス観測技術衛星2号
（いぶき2号, GOSAT-2）



温室効果ガス・水循環観測技術衛星（GOSAT-GW）で観測される二酸化炭素カラム平均濃度の想定図
（軌道考慮なし）

JAXA降水観測ミッションとの連携

- 研究成果のNEXRAシステムへの反映
  富岳でのNICAM-LETKF高速・省メモリ化

NIES GOSATプロジェクトとの連携

- NICAM-LETKFを用いた高精度炭素収支解析システム構築に向けた開発

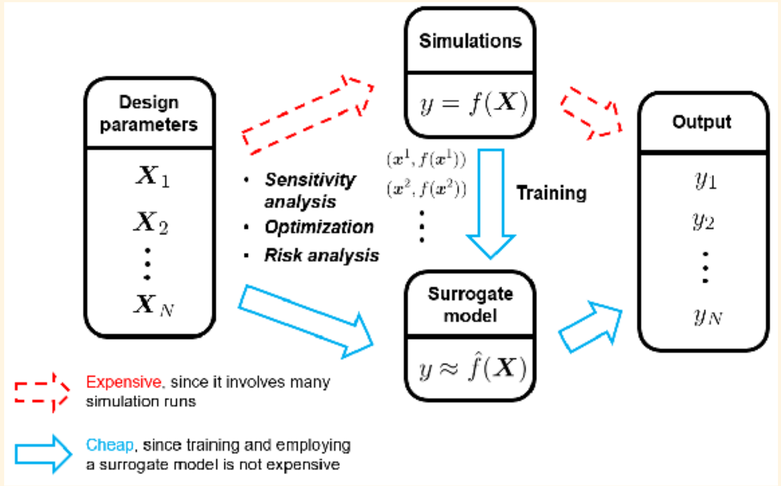
富岳のその先は？

ポスト「富岳」に向けた動き

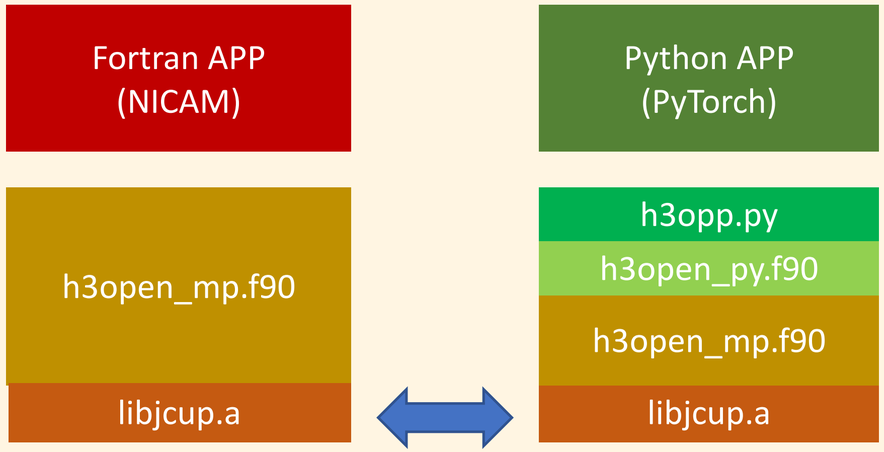
- NGACI (Next-Generation Advanced Computing Infrastructure) White Paper
  - 2020.11初版制定（https://sites.google.com/view/ngaci）
  - 2028年のシステム性能予測
    - メニーコア型: 50MWの電力消費でも最大1.8EFLOPS（富岳の3倍強）
    - GPU型：50MWで最大18EFLOPS（富岳の30倍強）
    - どちらのケースもメモリ性能は富岳の4倍、メモリ容量は富岳の2倍程度に留まる
      （ただし、大容量の不揮発性メモリストレージをその外に持つだろう）
  - 気象・気候分野にとってどういうことか？
    - メモリ性能に頼るアルゴリズムでは10年経っても大して速くしてもらえる見込みがない
    - フラッグシップ以外の調達でも、CPUのみのマシンは今よりもっと割高に。電気代も増加

次の100倍を狙うには？

- 物理モデルと高速代替モデルの併用
  - プロダクションランは機械学習によるいわゆるSurrogate Modelを利用する他に、
    向こう10年での劇的な高速化は見込めないだろうと考える（私見）
  - 既存の物理モデルの継続的な開発を止めず、段階的な適用を進める上でも現実的
- 科研費基盤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（2019-2024, PI:東大中島研吾先生）」での取り組み
  - 連成カプラーを用いた機械学習の「プラグイン」：異なる解像度間でもon-the-flyで学習データを収集



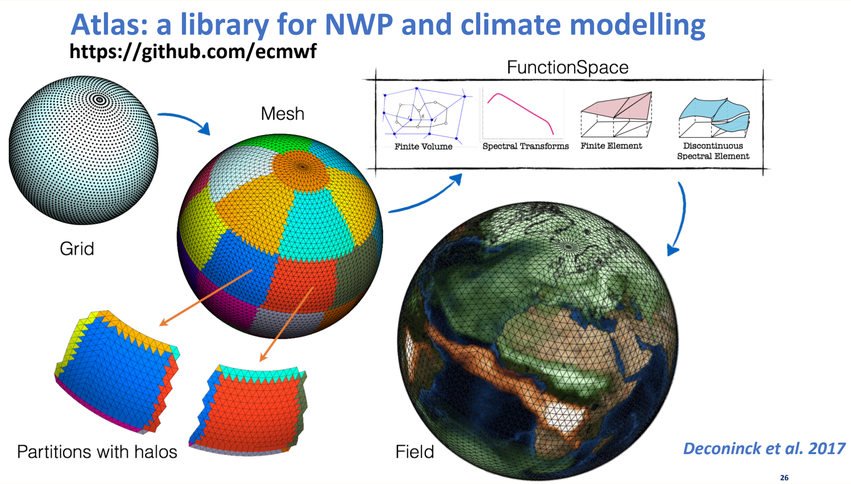
by Shuai Guo https://towardsdatascience.com/

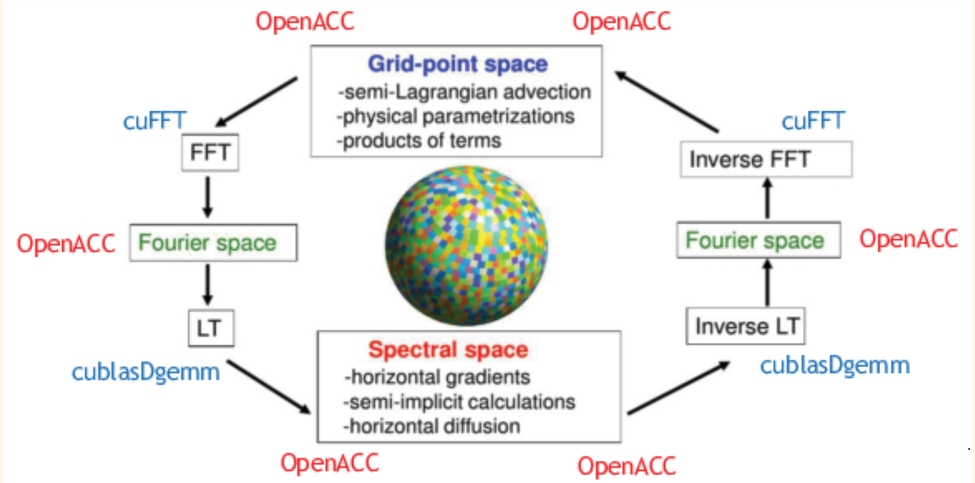


Providing teacher data on the fly

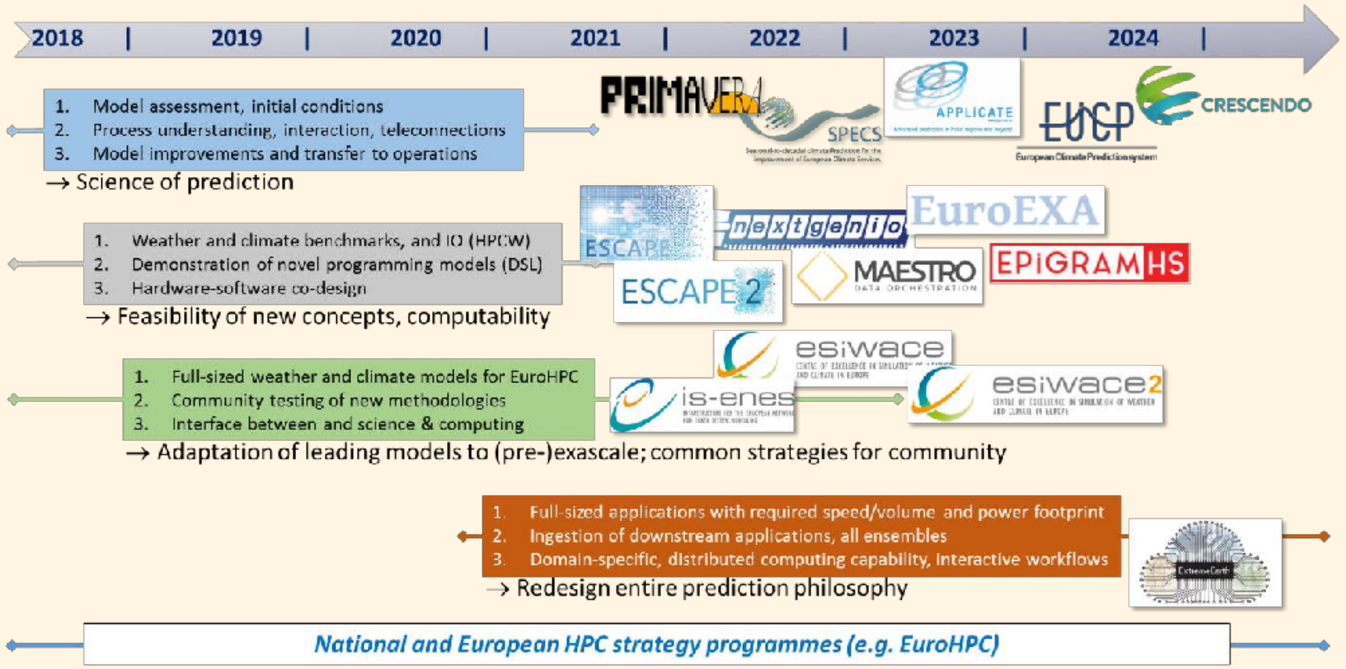
海外の動き

- 海外の多くの気象機関・研究コミュニティは、予測精度の向上のために水平数kmまで高解像度化していく流れは避けて通れないと考えている
  - 気象も、気候も
  - 高解像度化に必要な要素は最新のスパコンであり、最新のスパコンの変化に追随するにはプログラミングモデルを一新することも厭わない姿勢
    - ECMWFのAtlas, UFS(FV3GFS)のGT4Py, E3SMのKokkos, 等
  - AIを活用したモデル高速化も随所に





海外の動き（欧州）

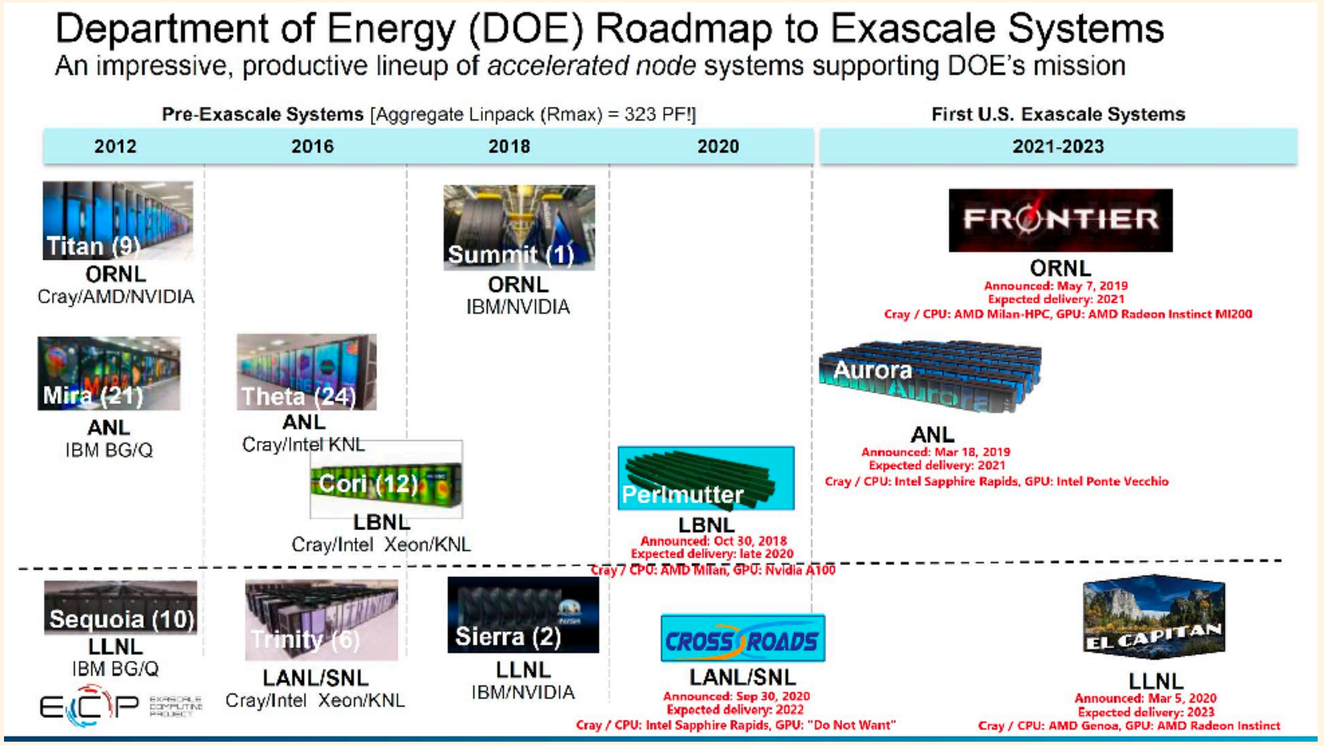


- EuroHPC（計算基盤）
- ESCAPE2（モデリング）
- ESiWACE2（Production）
- ExtremeEarth（新機軸）

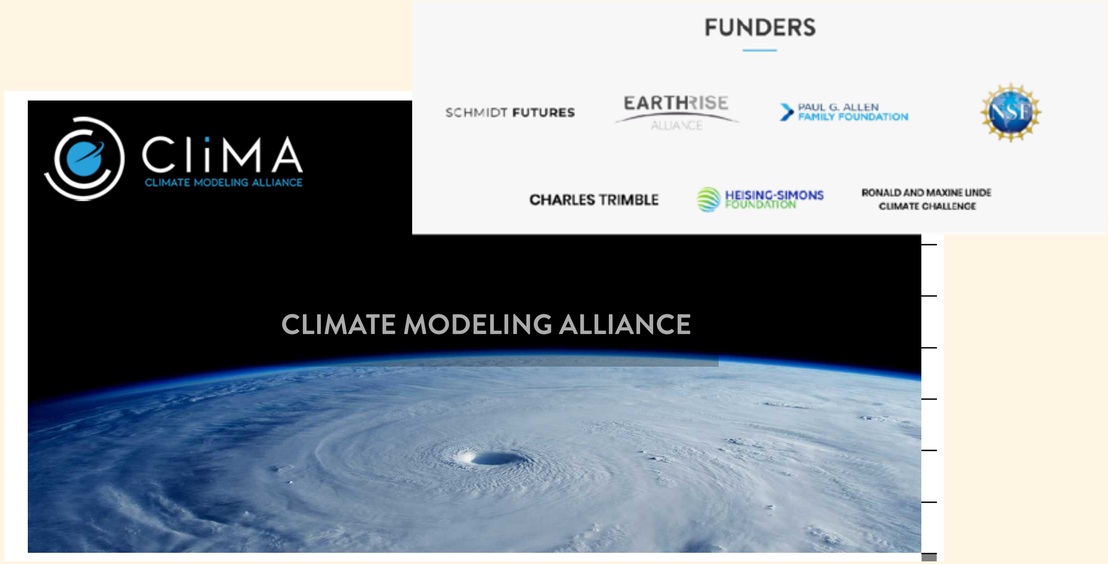
ECMWF: 30-40PFLOPS?, 2021-
MetOffice: 80PFLOPS?, 2022-
MeteoSwiss: CSCSにexa級マシン, 2023-

- Pre-exascale
  - LUMI (フィンランド, AMD CPU+GPU)
  - LEONARDO (イタリア, AMD CPU+NVIDIA GPU)
  - MareNostrum5 (スペイン，ヘテロシステム?)
- Petascale
  - ルクセンブルク、スロベニア、チェコ、ブルガリア、ポルトガル

海外の動き（米国）



NOAA: 24PFLOPS, 2022-



VULCAN

まとめに代えて

- 富岳プロジェクトは気象・気候アプリケーションにとっては大成功だった
  - メモリ律速のアプリケーションを救うマシン
  - 大胆しかし段階的なアプリケーションの改良ができた

- しかし、気象・気候シミュレーションの未来は明るいとはいえない
  - これまでの延長線上のマシンでは、ただ待っていても出来ることは増えない
  - 再び大胆な方針を打ち出す必要がある
  - 引き続きスパコン開発とタッグを組んで進める必要がある