HPC+AI for Earth Sciences at NERSC

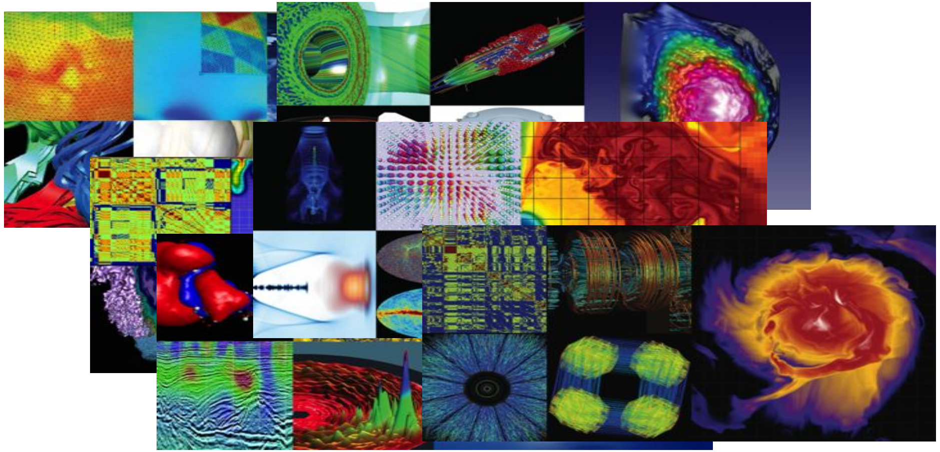
NERSC

Peter Harrington
Data & Analytics Services, NERSC
Lawrence Berkeley National Laboratory

Outline

- AI4Earth @ NERSC overview
- Data-driven forecasting state of practice
- FourCastNet background
- Scaling FourCastNet
- Entering the climate realm

NERSC: Mission HPC for the Dept. of Energy Office of Science



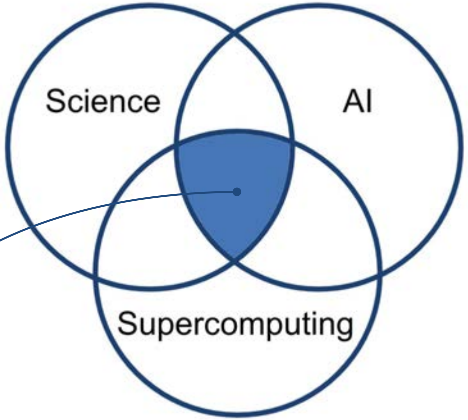
Large compute and data systems

- Perlmutter: ~7k A100 GPUs
- 30 PB all-flash scratch filesystem
- 128PB Community Filesystem ....

Broad science user base

- 7,000 users,
- 800 projects,
- 700 codes

NERSC AI Strategy

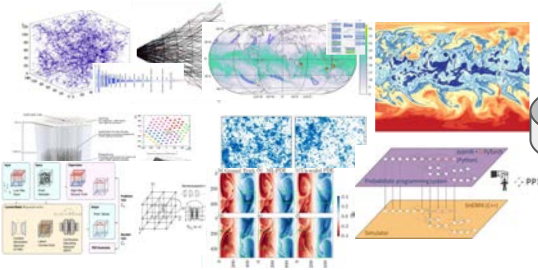


- The **intersection** of HPC, AI, & science

- Focus activities in three main areas:

NERSC AI

Methods and Applications



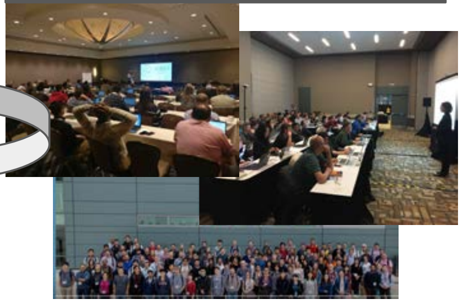
Deployment

Automation | **Interactivity**

Software Frameworks and Libraries

Systems w/ Accelerators

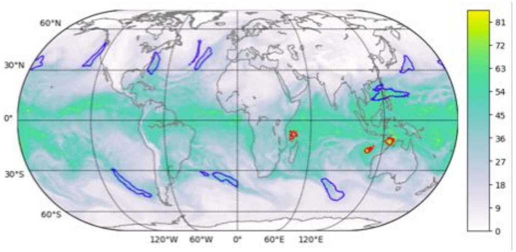
Empowerment



NERSC AI in Earth Sciences

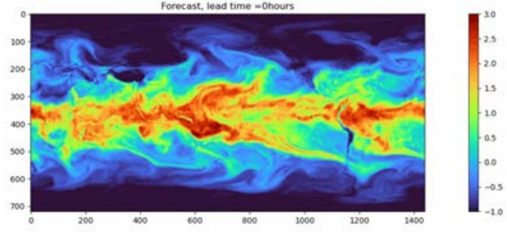
We see cutting-edge earth science AI pushing supercomputers to their limits

- AI beginning to **revolutionize weather & climate modeling**
- These challenges **drive innovation in both ML and HPC**



Highlighting some example projects:

- 2018 DeepCAM (Gordon Bell Prize)
  - High-resolution climate segmentation model
  - Detection of hurricanes, atmospheric rivers, tropical storms
  - Aid and accelerate climate analytics
- 2022 FourCastNet
  - High-resolution forecasting model
  - Superior global medium-range weather skill



FourCastNet: Data-driven atmospheric forecasting at scale

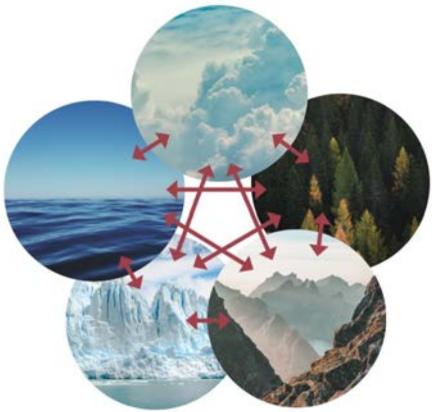
NERSC

BERKELEY LAB

U.S. DEPARTMENT OF ENERGY | Office of Science

Weather forecasting

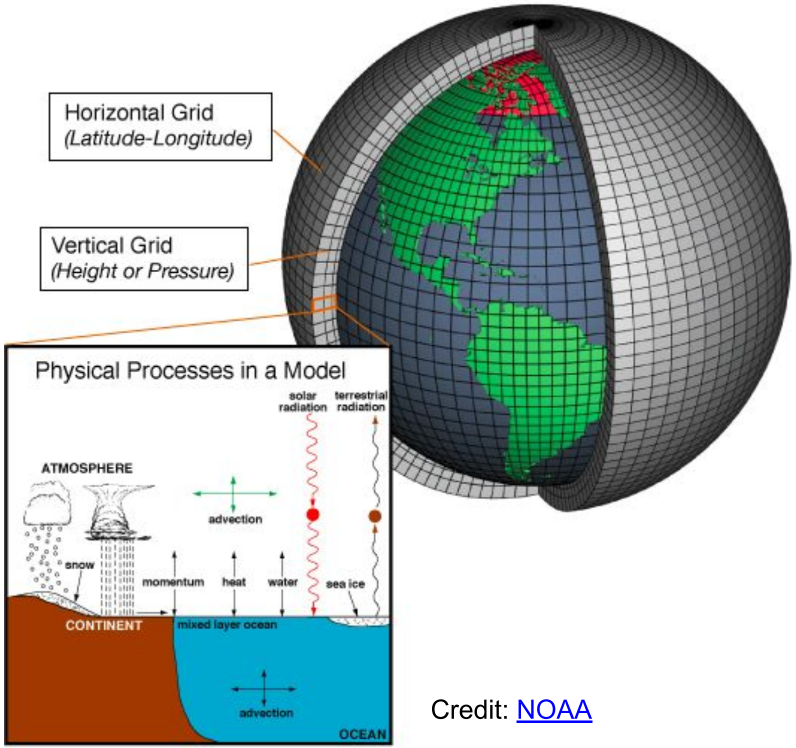
- Urgency under climate change:
    - Reducing uncertainties, predicting extremes, disaster mitigation, etc…
- Entire earth system: complex phenomena across wide range of physical scales



- Traditional NWP consumes substantial HPC resources
    - Dedicated machines running physical models + data assimilation 24/7

Ingredients of NWP

- PDEs w/ hundreds of variables, complex parameterizations for subgrid and multi–physics processes
- Large ensembles for UQ and long-term forecasting
- **High resolution grids** (computational cost scales as ~fourth power)

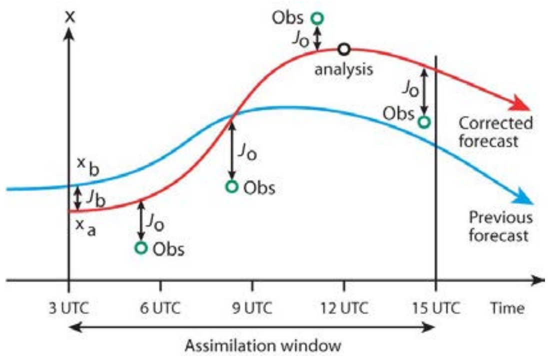


Credit: NOAA

Ingredients of NWP

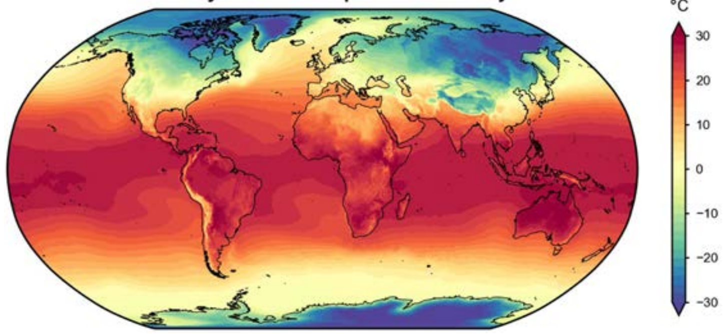
- Data assimilation of observations on (6-12)hr cycles (4D-Var EnKF for ECMWF IFS model)
    - Operationally, this produces “analysis” states

- Additional post-processing: “reanalysis”
    - ERA5 current best for atmospheric reanalysis data



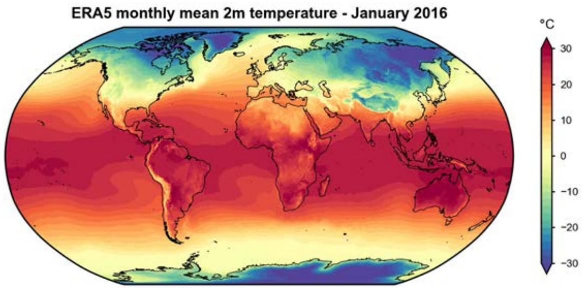
Credit: [ECMWF]

ERA5 monthly mean 2m temperature - January 2016

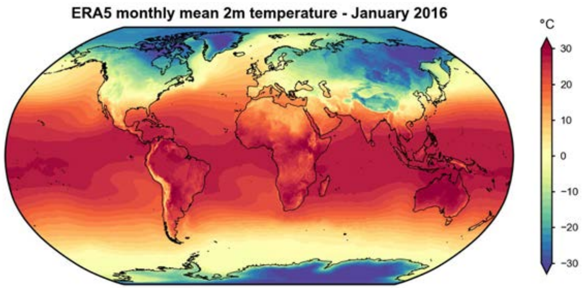


Data-driven weather & climate forecasting

- ERA5 reanalysis dataset: 40 years, 25km global grid, assimilated with observations. “Best available estimate of earth’s atmospheric state”



Data-driven weather & climate forecasting

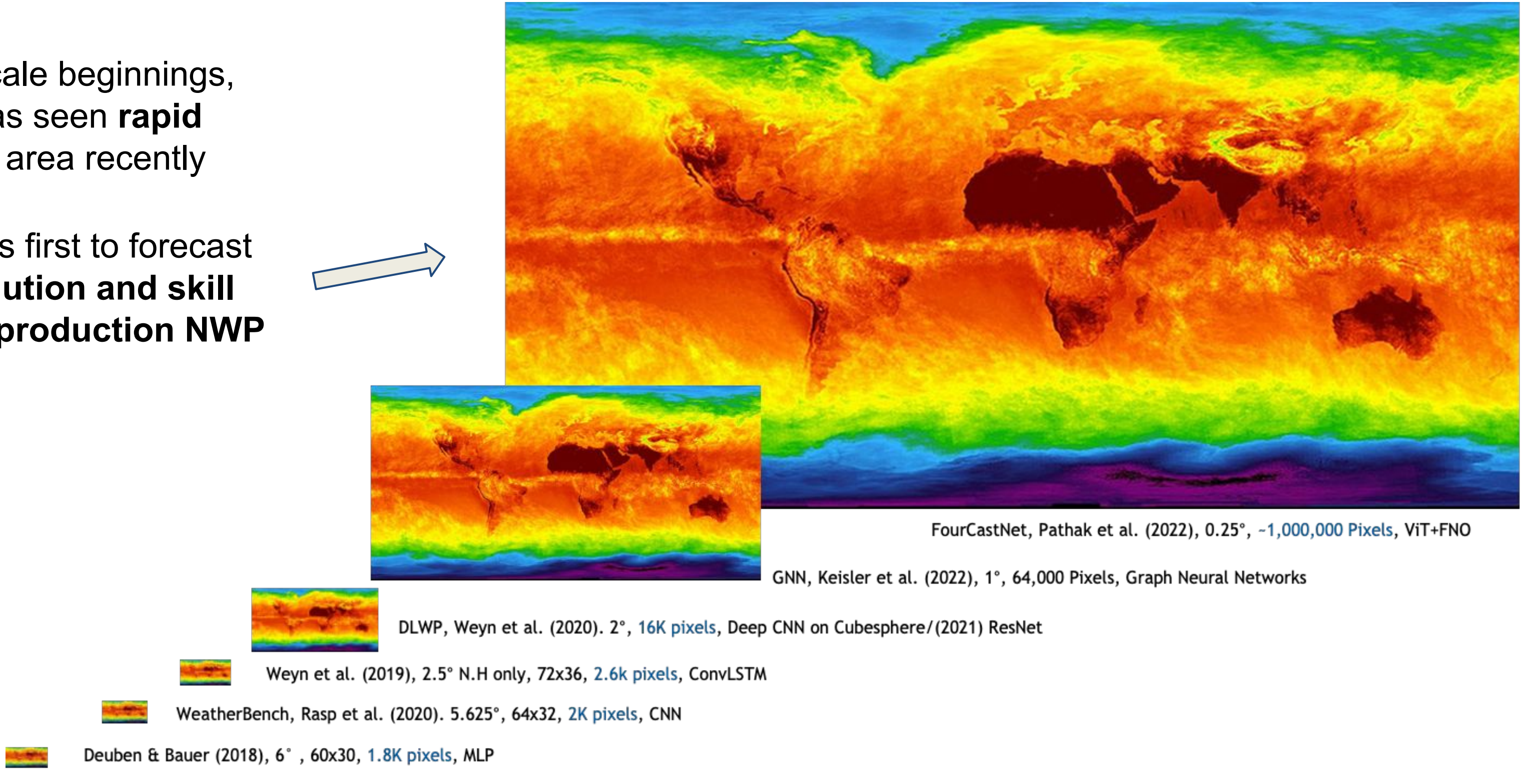


- ERA5 reanalysis dataset: 40 years, 25km global grid, assimilated with observations. "Best available estimate of earth's atmospheric state"
- Potential gold mine for data-driven forecasting models:
    - Overcome model biases; use mixed-precision & GPUs for fast inference
    - Challenge: **very high-dimensional** state space (~150M 'pixels' for a single variable)

Data-driven weather & climate forecasting

- Despite small-scale beginnings, deep learning has seen **rapid progress** in this area recently

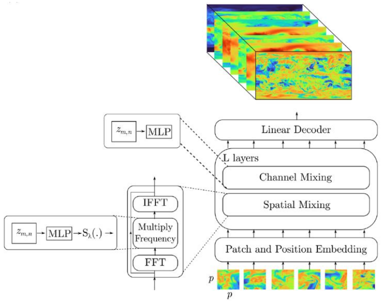
- FourCastNet was first to forecast weather at **resolution and skill comparable to production NWP models like IFS**



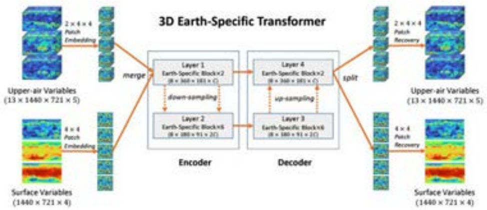
Data-driven weather & climate forecasting

- Now, **multiple deep learning models approach or outperform forecast skill** of traditional NWP

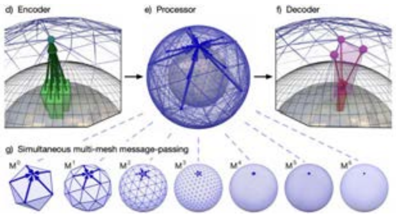
- Training recipes vary, but general theme is training large networks on a large chunk of ERA5
    - All (except GraphCast) are transformer-based
    - Large scale: some models require **weeks of training on hundreds of GPUs**



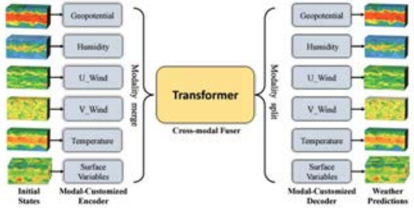
FourCastNet



Pangu-Weather



GraphCast



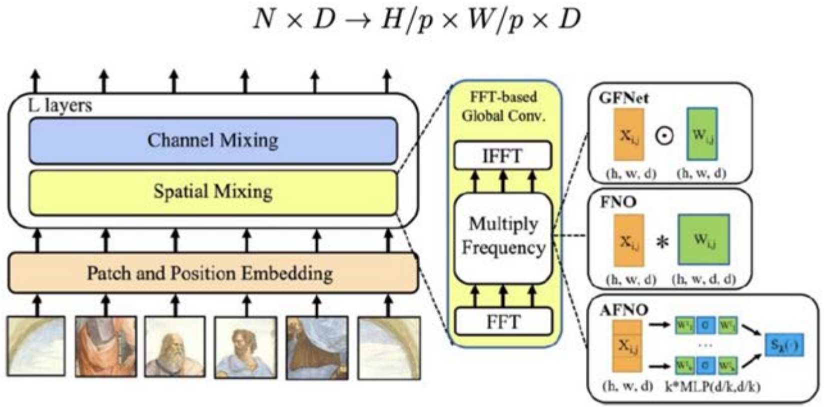
FengWu

Scaling FourCastNet on HPC systems

NERSC BERKELEY LAB U.S. DEPARTMENT OF ENERGY Office of Science

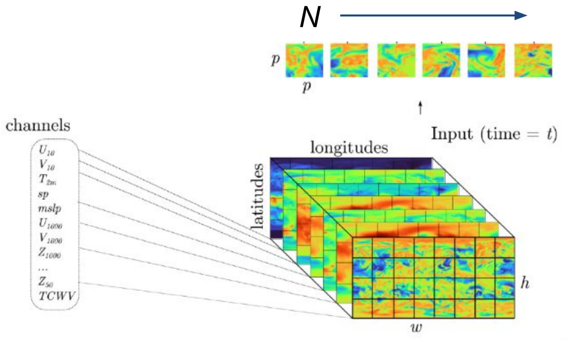
FourCastNet model: Adaptive FNO

Adaptive FNO architecture replaces expensive self-attention in vision transformers with token mixing via FFTs



Models	Complexity (FLOPs)	Parameter Count
Self-Attention	$N^2d + 3Nd^2$	$3d^2$
AFNO (ours)	$Nd^2/k + Nd\log N$	$(1+4/k)d^2 + 4d$

N is set by **resolution and patch size**



Guibas et al. "Adaptive Fourier Neural Operators: Efficient Token Mixing for Transformers." arXiv:2111.13587 (ICLR 2022).

FourCastNet: scaling for future

- **Resolution is key**: future efforts will leverage data at 10km, 5km, (eventually) 1km grids. How do we scale up?

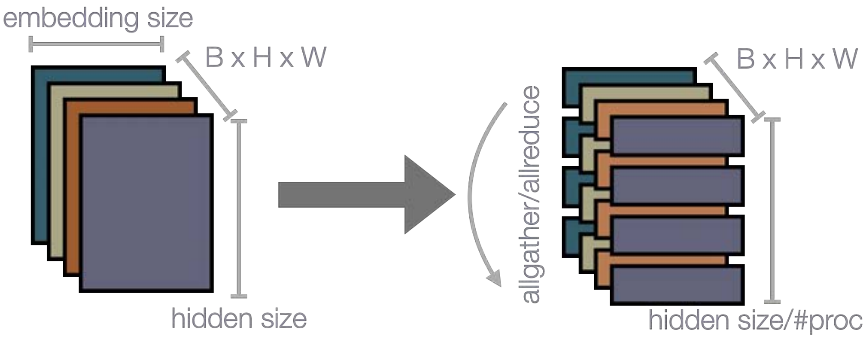
Models	Complexity (FLOPs)	Parameter Count
AFNO (ours)	$Nd^2/k + Nd\log N$	$(1+4/k)d^2 + 4d$

	Current (25 km)	Intermediate (5 km)	Large (1 km)
N (p = 1)	1M	25M	625M
FFTs	720 x 1440 (d of them)	3600 x 7200 (d of them)	18k x 36k (d of them)
Matmul	[4d x d] * [d] (N of them)	[4d x d] * [d] (N of them)	[4d x d] * [d] (N of them)

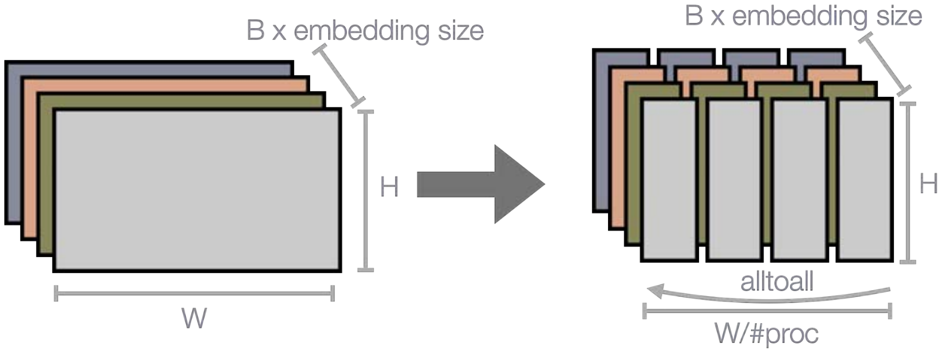
- Different parallelization strategies needed to balance spatiotemporal resolution, model capacity, and hardware constraints

Beyond data parallel training

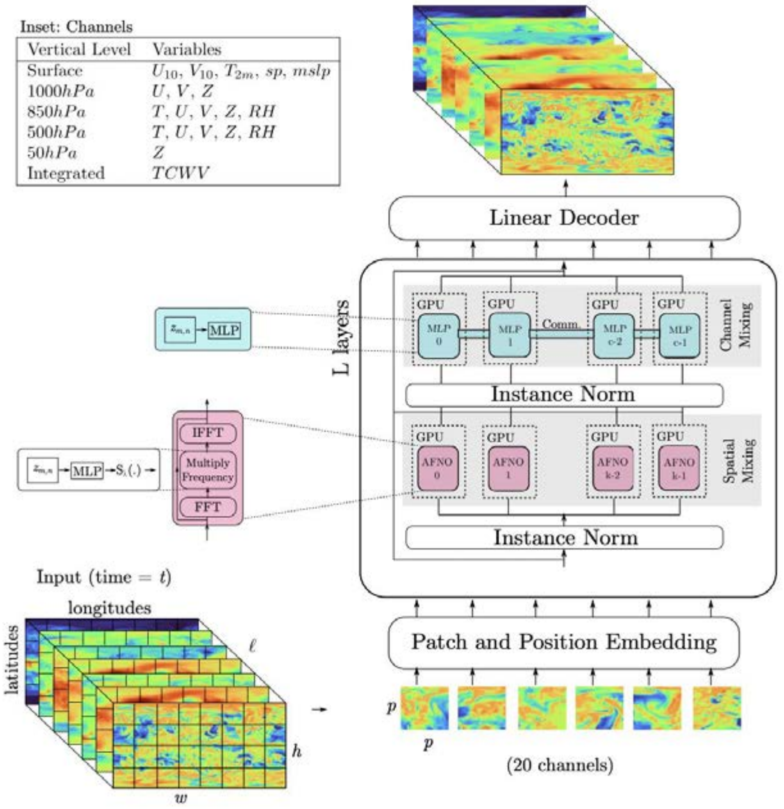
- **Feature parallelization**
  splitting channel dim, dense layers become distributed matrix multiplications



- **Domain decomposition**
  splitting height and width, FFT become distributed, LayerNorm needs to exchange stats

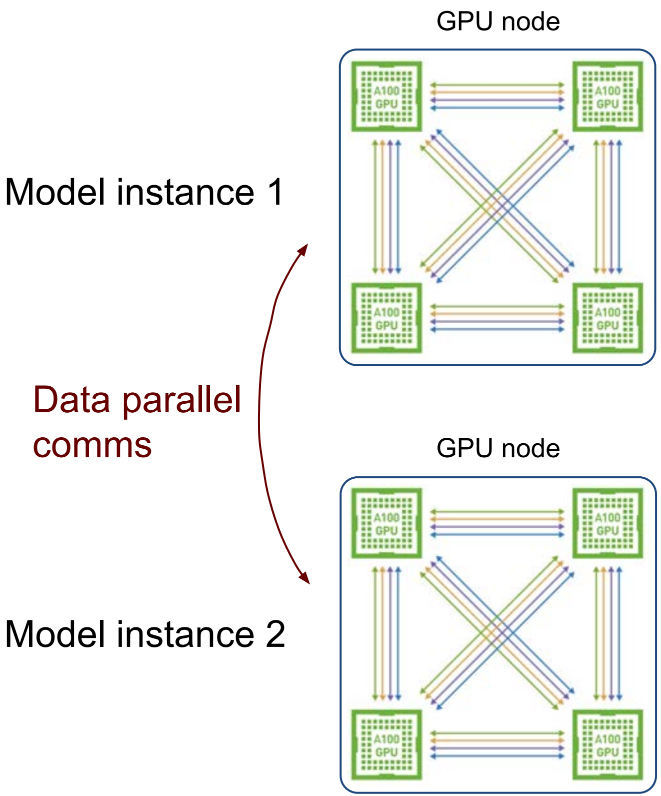


FourCastNet++: hybrid data-model parallel training



- Split MLP features along channel dimensions for model parallelism
- FFT-based spatial mixing operates on disjoint blocks, so embarrassingly parallel

FourCastNet++: hybrid data-model parallel training



- Model parallelism across GPUs within a node: model instance
    - Exploit high-speed **NVLink** conneciton

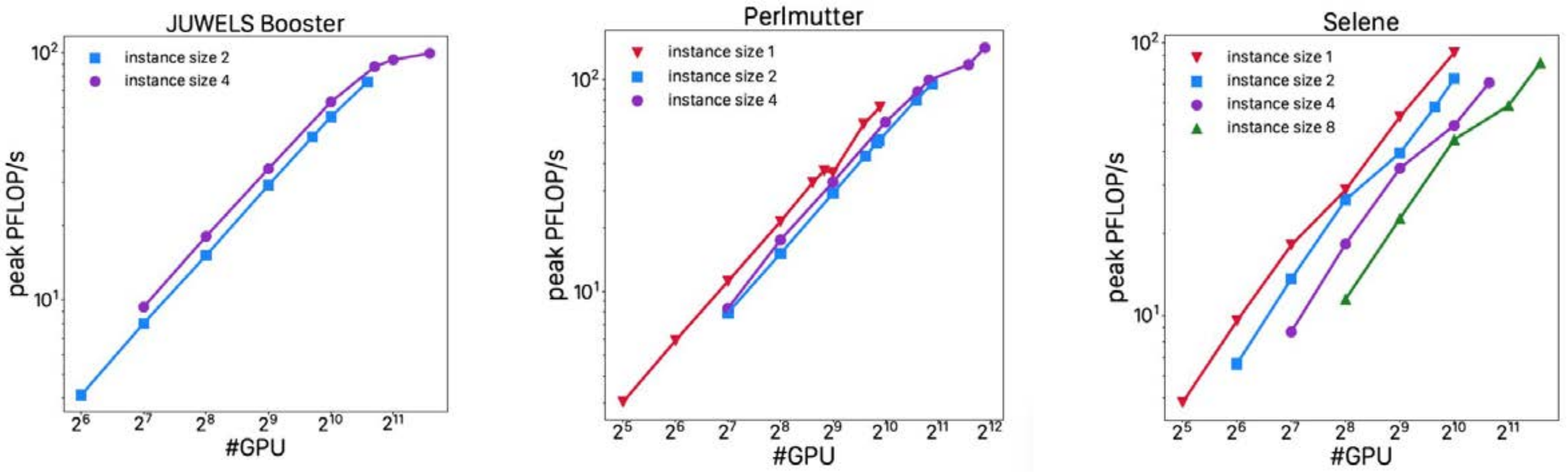
- Data parallelism across model instances
    - Gradient reduction during SGD happens across model instances

- Implement with separate communicators for model and data parallel comms

- Similar to state-of-the-art massive LLMs, e.g. Megatron-LM

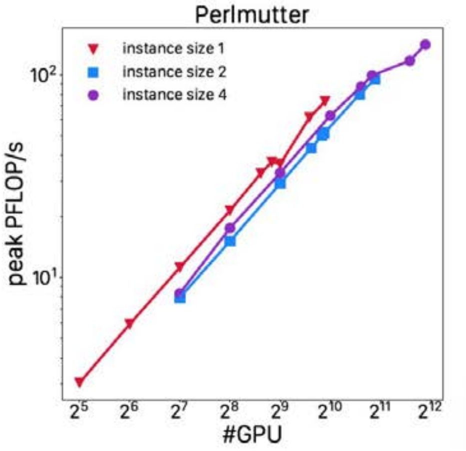
Scaling results

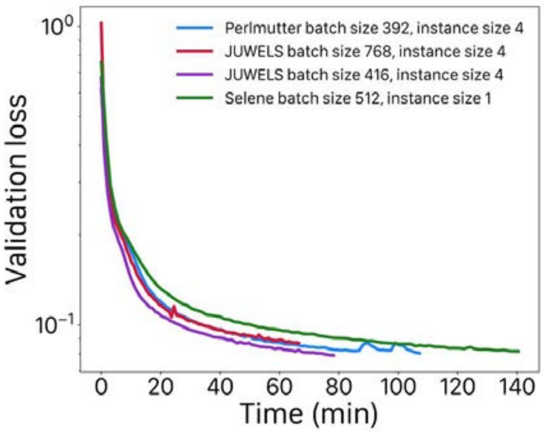
- Preliminary scaling study conducted on 25km resolution, with patch size=4
- Scaling results across three HPC systems: **140.8 PFLOP/s on ~4000 GPUs**



Scaling results

- Preliminary scaling study conducted on 25km resolution, with patch size=4
- Scaling results across three HPC systems: **140.8 PFLOP/s on ~4000 GPUs**
- Time to solution can be reduced to ~1hr or less



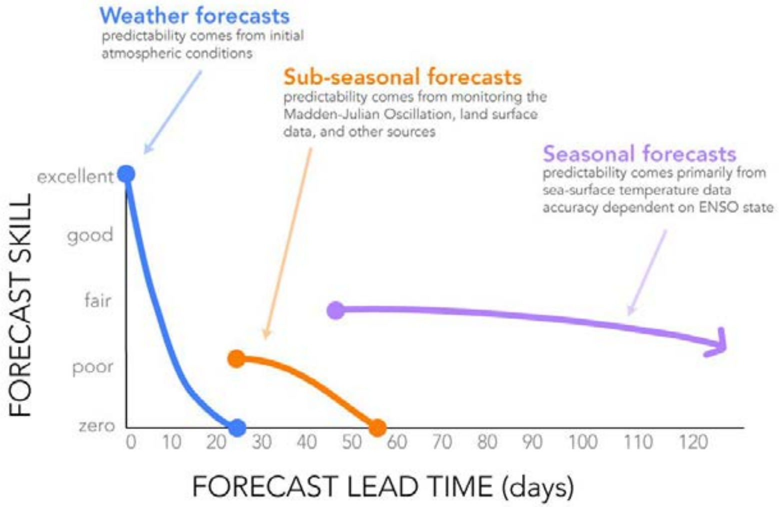


Details:
Kurth et al.
PASC 2023 Best Paper
https://dl.acm.org/doi/abs/10.1145/3592979.3593412

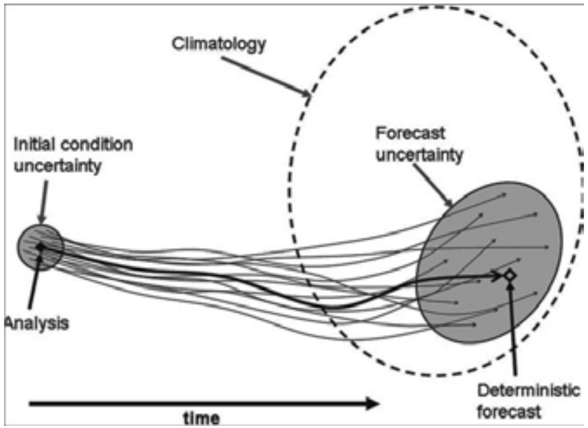
Looking forward: from weather to climate

From weather to climate

- Deep learning offers exciting new capabilities for medium-range weather forecasting
- What can we do as we move beyond that?



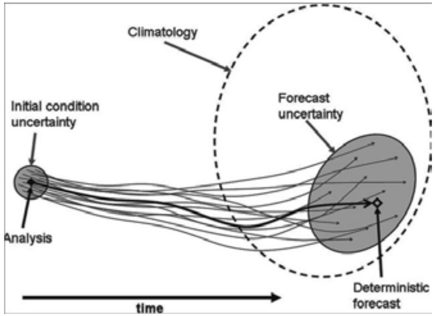
https://iri.columbia.edu/news/qa-subseasonal-prediction-project

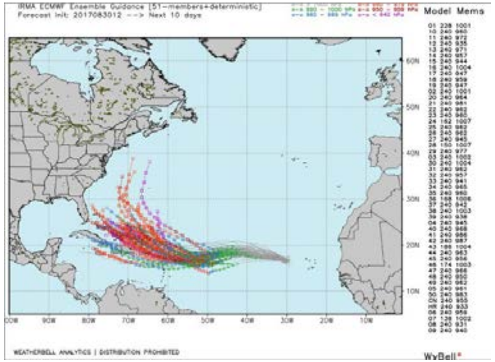


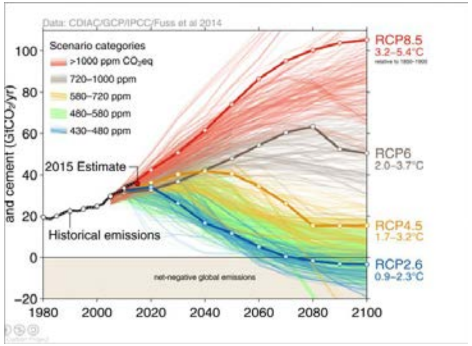
From weather to climate

- Fast inference enables (very) large-scale ensembles

- **Hindcasting** with large ensembles can be done against both historical record as well as weather states from future climate scenarios: **characterize uncertainty/likelihoods**



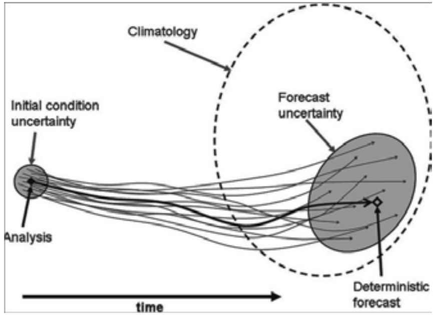


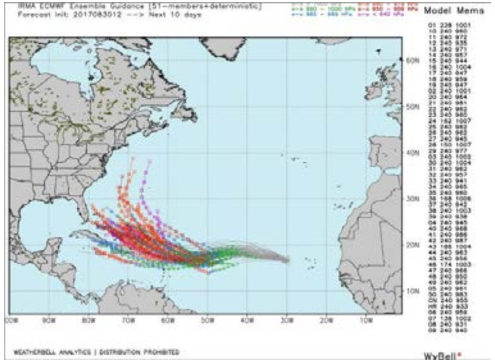


From weather to climate

- Characterizing likelihood of extreme events is a major challenge of climate science

- We are now positioned to do so with AI+HPC:
  - Couple fast forecasting models with advanced climate analytics engines like TECA
  - Answer questions like “how does the frequency of hurricanes look in 2050” **with error bars**





Other future directions

- **Computational scale**
  - train with 10km, 5km, 1km data
  - All vertical levels for full 3D structure
  - "Full" spatiotemporal model
- **Explicitly incorporating physics**
  - Conservation terms, better stability, etc.
  - Coupling with climate models
- **Better ensembling**
  - Generative models, learned perturbations
  - Uncertainty quantification
- **Direct data assimilation into DL model**