Machine Learning Research for Climate Change and Environmental Sustainability

Claire Monteleoni

INRIA Paris

University of Colorado Boulder

Inria

December 2021: Boulder County, Colorado

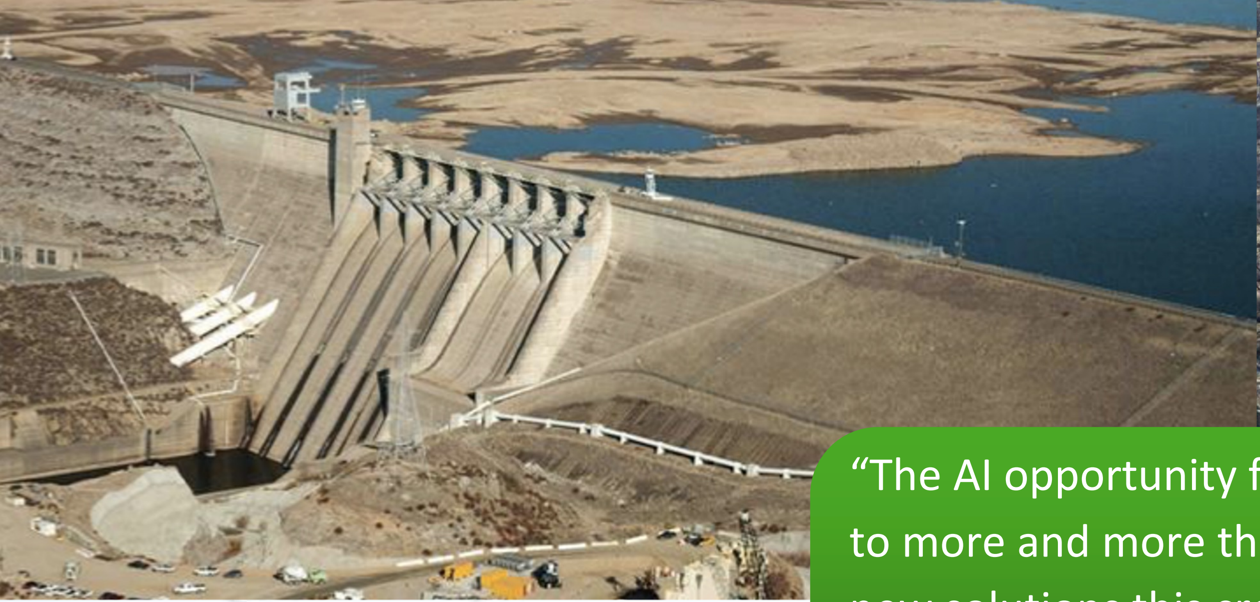
- Snow drought conditions through fall and winter 2021 created dry land-cover
- 80-100 mph winds, combined with ignition, launched an uncontrollable “fire storm”
- Loss of 2 lives. 1000 homes and 20 businesses were destroyed, and more damaged

January 2018: Montecito, Santa Barbara County

- Thomas Fire destroyed 1063 structures and led to poor air quality
- Intense rainfall as the fire was nearing containment produced a debris flow
- 23 lives and over 130 homes were lost
- Damage to critical transportation and water resource infrastructure




Machine learning can shed light on climate change



“The AI opportunity for the Earth is significant. Today’s AI explosion will see us add AI to more and more things every year.... As we think about the gains, efficiencies and new solutions this creates for nations, business and for everyday life, we must also think about how to maximize the gains for society and our environment at large.”

– The World Economic Forum: Harnessing Artificial Intelligence for the Earth. 2018

Climate Informatics is based on the vision that Machine learning can shed light on climate change

2008 Started research on Climate Informatics, with Gavin Schmidt, NASA

2010 "Tracking Climate Models" [Monteleoni et al., NASA CIDU, Best Application Paper Award]

2011 Launched International Workshop on Climate Informatics, New York Academy of Sciences

2012 Climate Informatics Workshop held at NCAR, Boulder, for next 7 years

2013 "Climate Informatics" book chapter [M et al., SAM]

2014 "Climate Change: Challenges for Machine Learning," [M & Banerjee, NeurIPS Tutorial]

2015 Launched Climate Informatics Hackathon, Paris and Boulder

2018 World Economic Forum recognizes Climate Informatics as key priority

2021 Computing Research for the Climate Crisis [Bliss, Bradley @ M, CCC white paper]

2022 First batch of articles published in Environmental Data Science, Cambridge University Press

2023 12th Conference on Climate Informatics and 9th Hackathon, Cambridge, UK

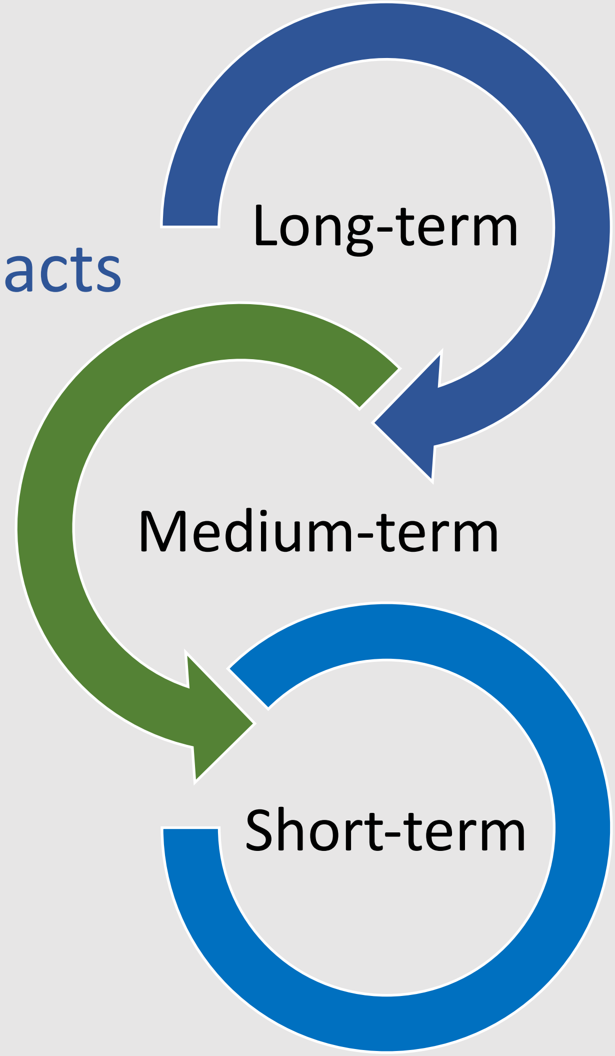
2024 13th Conference on Climate Informatics and 10th Hackathon, Turing Institute, London, APRIL

AI Research for Climate Change and Environmental Sustainability

- Machine Learning for Climate Science
  Understanding and Predicting Climate Change and Impacts

- Machine Learning for Climate Mitigation
  Accelerating the Green Transition

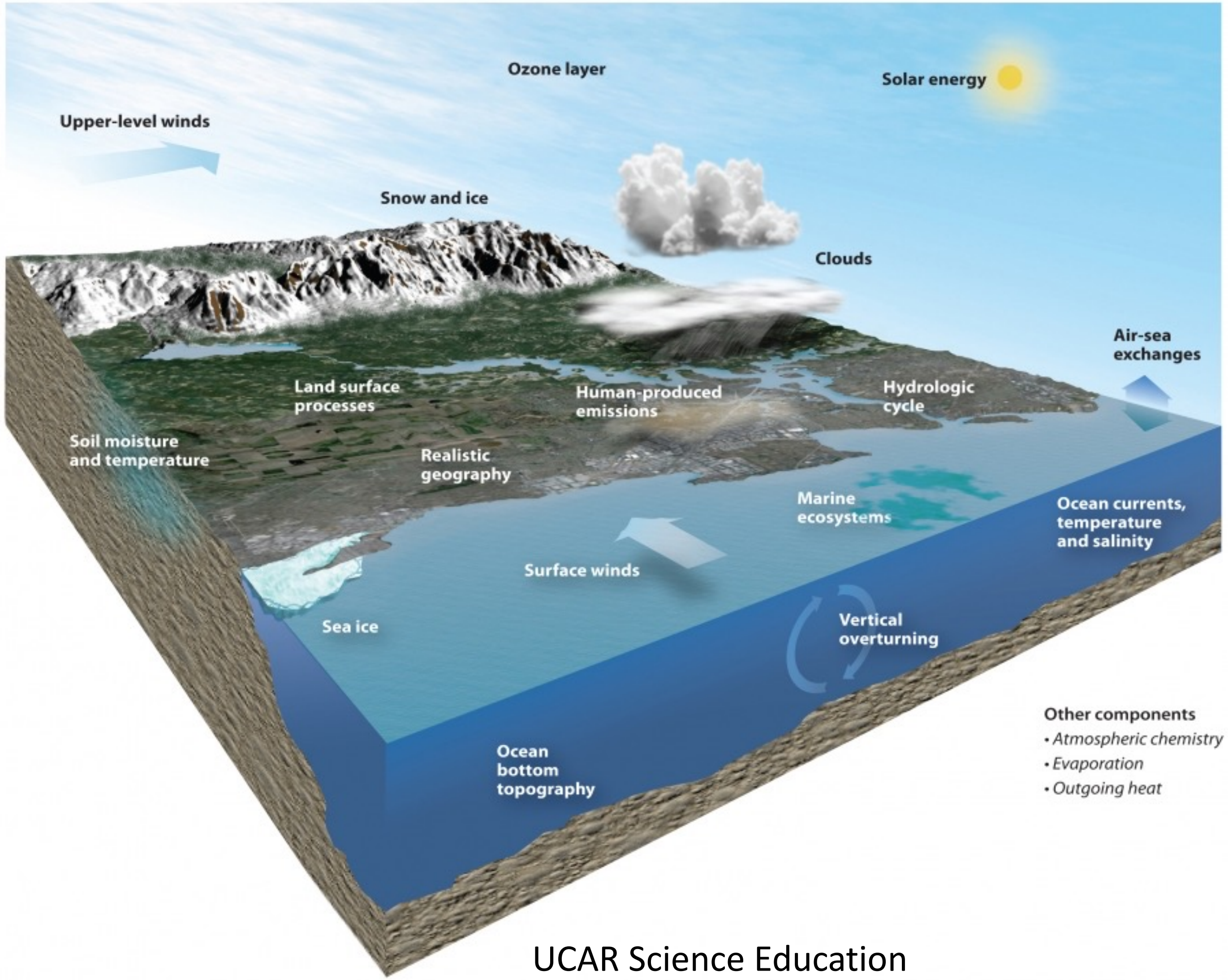
- Machine Learning for Climate Adaptation
  Extreme Weather and Cascading Hazards



Our Climate Informatics research also addresses open problems in Machine Learning

- Online learning with spatiotemporal non-stationarity
- Prediction at multiple timescales simultaneously
- Anomaly detection with limited supervision
- Tracking highly-deformable patterns

Machine Learning for Understanding and Predicting Climate Change



UCAR Science Education

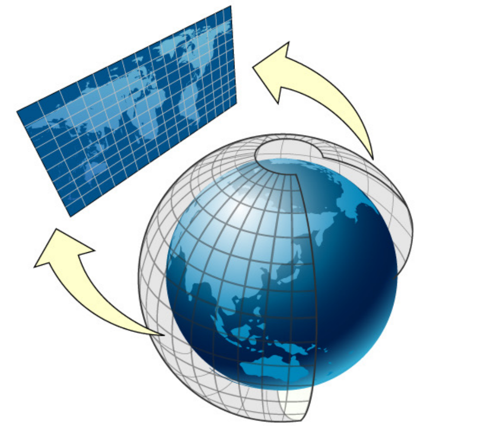
Robustify climate model ensemble forecasts using online learning for non-stationary spatiotemporal data [Multiple papers 2009-2020, e.g., AAAI 2012, ALT 2020]

Causal information hubs in Pacific ENSO region [Saha et al., Climate Informatics 2019]

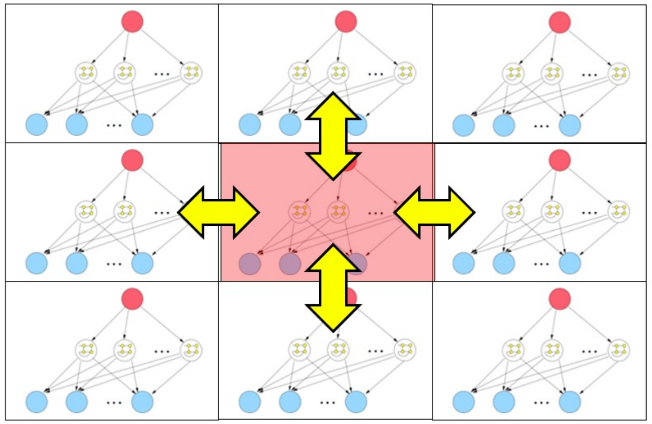
Project long-term sea-level rise:

- NASA project using satellite altimetry [Sinha et al., AGU 2022, ICLR 2023 workshop] CU & NCAR
- NSF i-HARP project on ice-sheet contributions to sea-level rise

Online learning with spatiotemporal non-stationarity



Learning when the target concept can vary over time, and multiple other dimensions (e.g., latitude, longitude)



We can exploit local structure in space and time

We can learn the level of non-stationarity in time and space

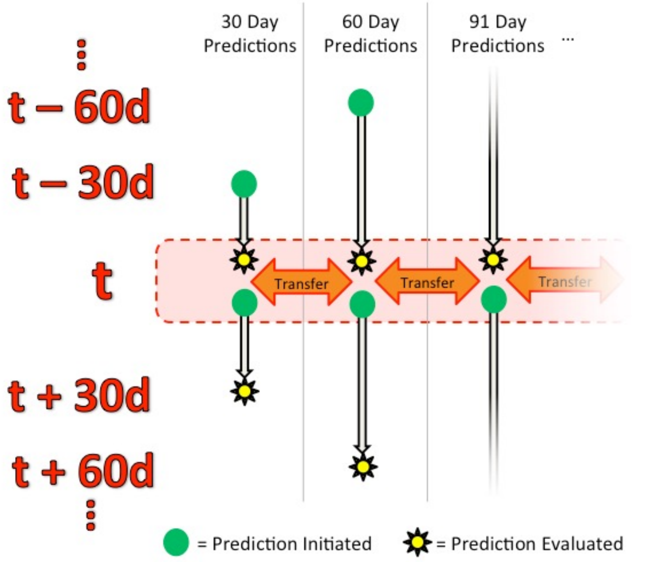
[McQuade and Monteleoni, AAAI 2012] extended [Monteleoni & Jaakkola, NeurIPS 2003; Monteleoni et al. SAM 2011] to multiple dimensions

This framework for online learning was open in machine learning

New “regret” framework: [Cesa-Bianchi, Cesari, & Monteleoni, ALT 2020]

Prediction at multiple timescales simultaneously

Applications to both climate science, and financial volatility: [McQuade and Monteleoni, CI 2015; SIGMOD DSMM 2016]



ML for the Green Transition

Accelerate the renewable energy transition

Week-ahead solar irradiance forecasting via deep sequence learning [Sinha et al., CI 2022] w/ NREL

ML to downscale climate model data for renewable energy planning:

- [Harilal et al., NeurIPS workshop 2022] Climate Change AI / Future Earth project w/ NREL, IIT-Roorkee
- INRIA Défis w/ Électricité de France

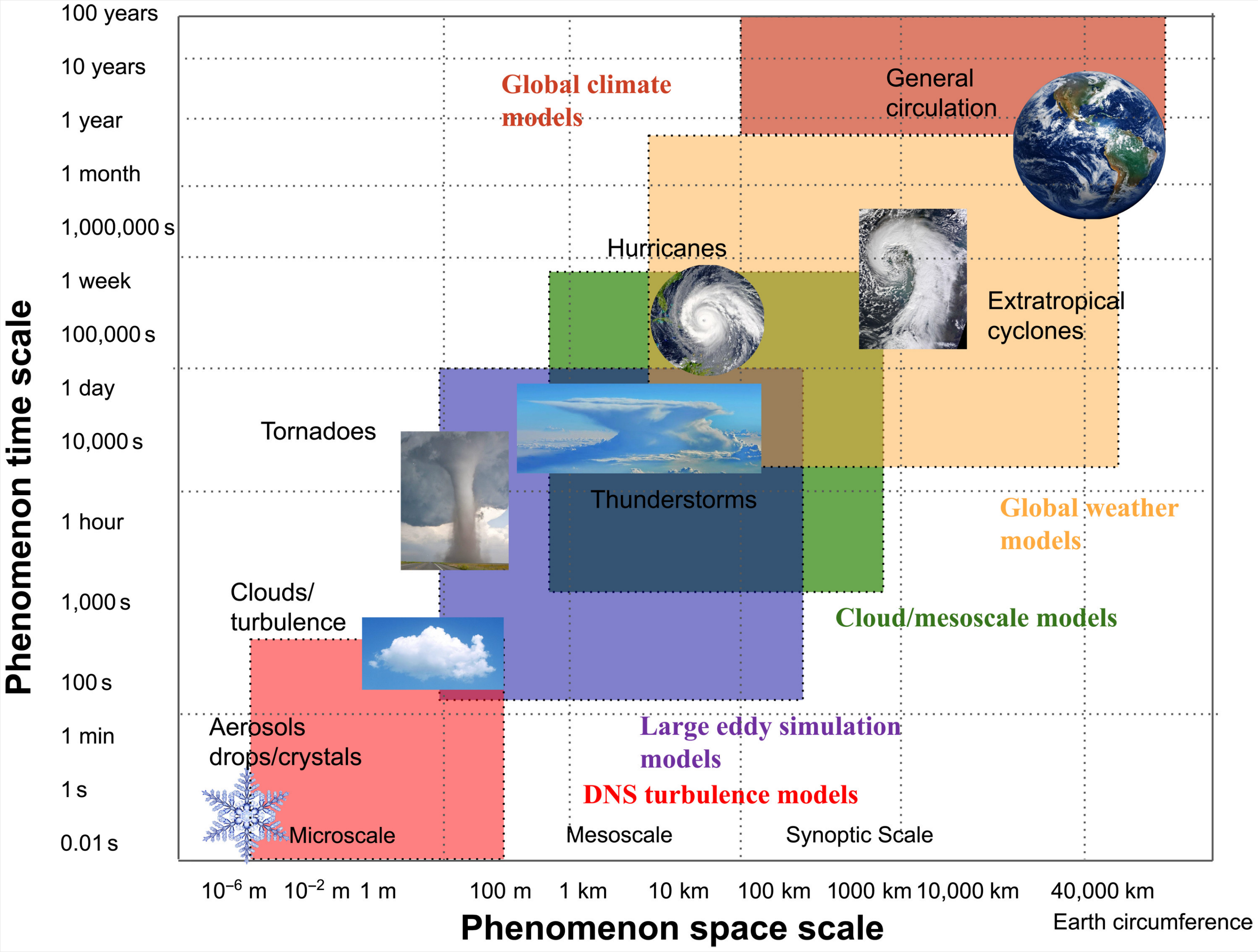
ML for modeling the effects of land-use change on CO_2 emissions

Downscaling climate model simulations

Global climate model simulations are coarser scale (in space and time) than needed for:

- Climate change mitigation
- Climate change adaptation

<u>Approach</u>: Use ML to downscale climate model data to relevant scales



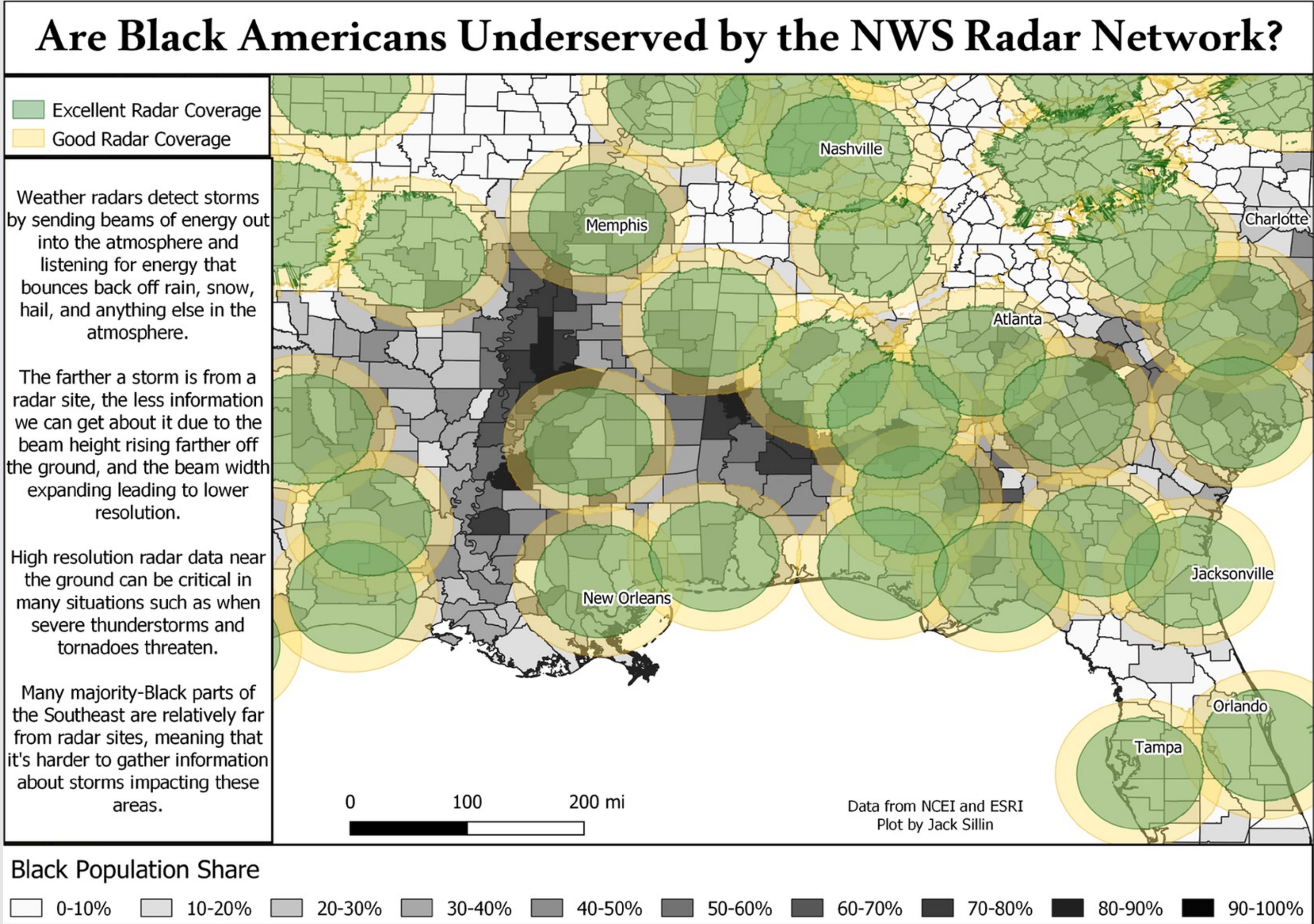
[Gettelman, et al., Science Advances, 2022]

Semi/Unsupervised downscaling: Equity motivation

- Train models in data-rich regions and apply them in data-poor regions
  - Can evaluate them against supervised learning models in data-rich regions
  - Can fine-tune them using the limited data in the data-poor regions

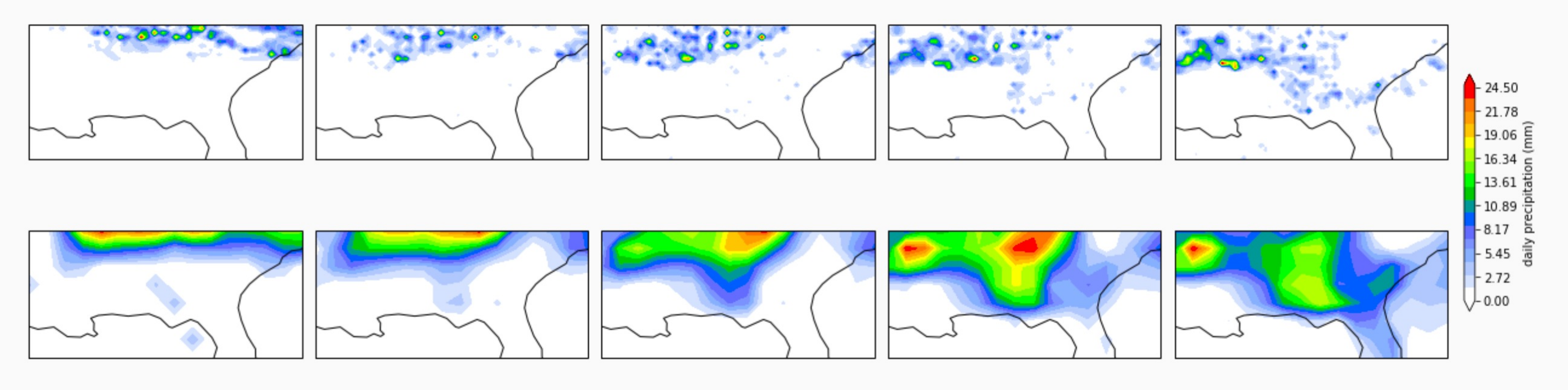
- Contribution to climate data equity
  - Global scales:
    - Global North historically emitted more carbon; Meanwhile there’s typically more data there
    - Global South is suffering the most severe effects of the resulting warming
  - Local scales (e.g. legacy of environmental injustice in USA)
    - Learn “virtual sensors”

"Many majority-Black parts of the Southeast [USA] are relatively far from radar sites, meaning that it's harder to gather information about storms impacting these areas."



Credit: Jack Sillin, in [McGovern et al., Environmental Data Science, 2022]

ClimAlign: Unsupervised, generative downscaling

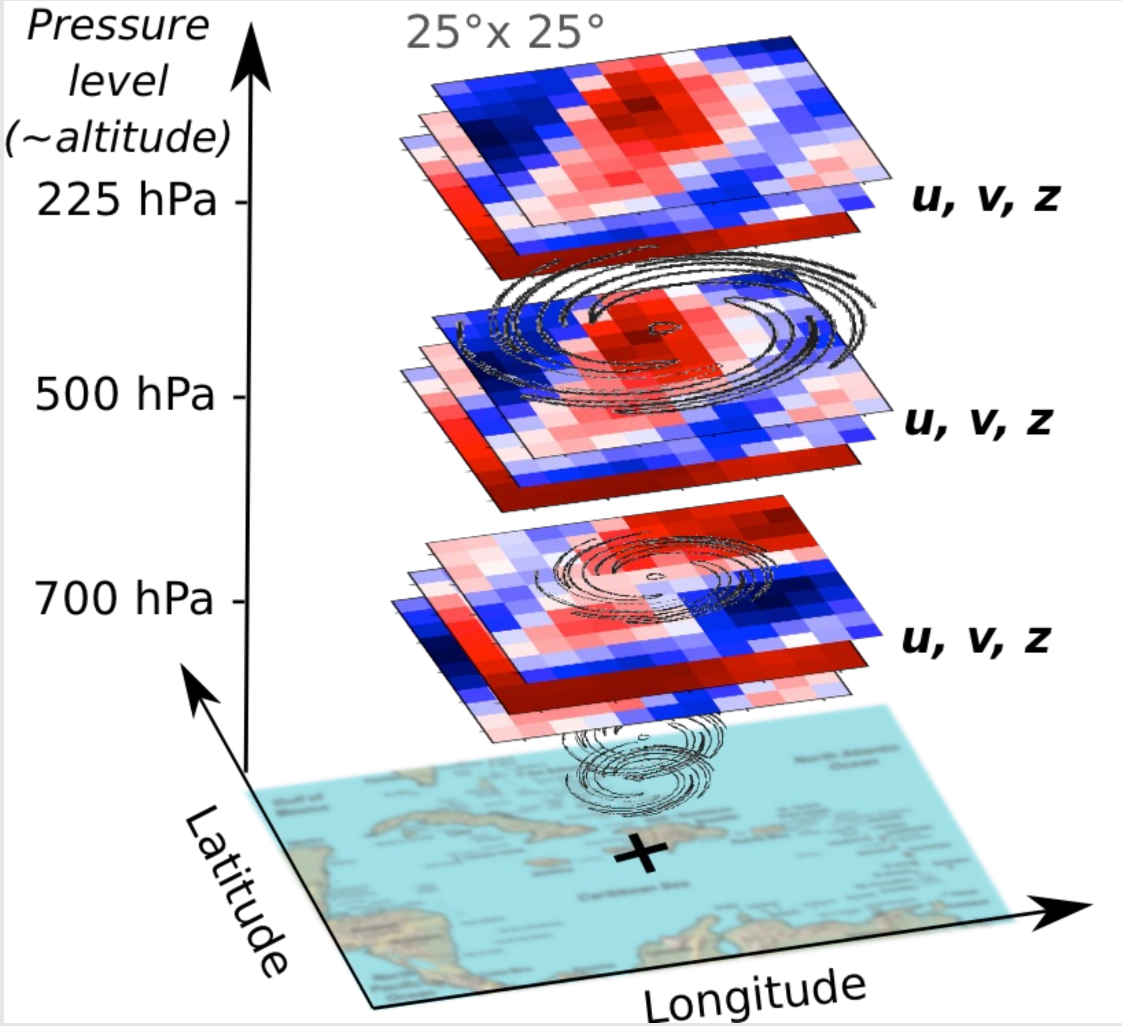


General downscaling technique via domain alignment with normalizing flows
[AlignFlow: Grover et al., AAAI 2020][Glow: Kingma & Dhariwal, NeurIPS 2018]

- **Unsupervised**: do not need paired maps at low and high resolution
- **Generative**: can sample from posterior over latent representation OR sample conditioned on a low (or high!) resolution map
- **Intepretable**, e.g., via interpolation

[Groenke, et al., Climate Informatics 2020]

Machine Learning for Extreme Weather and Cascading Hazards



[Giffard-Roisin et al., Frontiers 2020]

Hurricane track prediction via fused CNNs

[Giffard-Roisin et al., Climate Informatics 2018; Frontiers 2020]

Forecasting Indian Summer Monsoon precipitation extremes

[Saha et al. Climate Informatics 2019; 2020] w/ India Meteorological Department (IMD)

Avalanche detection using CNN; VAE

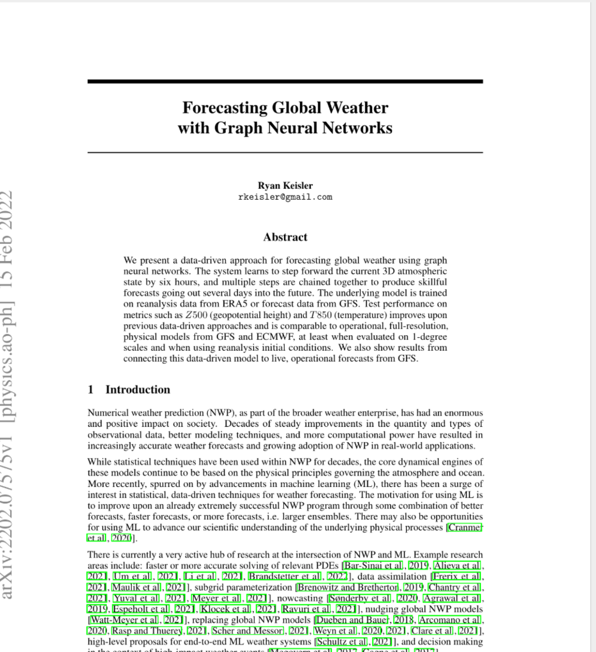
[Sinha et al., Climate Informatics 2019; 2020] w/ Météo-France

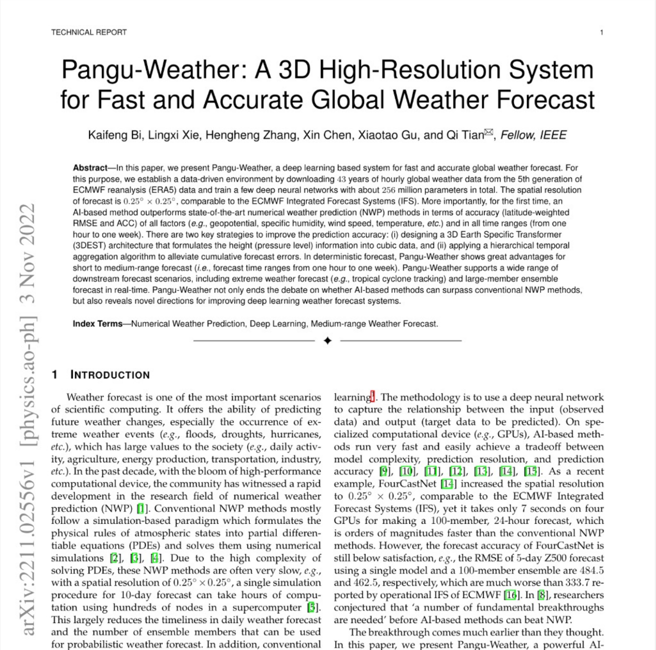
How to use “ML foundation models for weather” to forecast extreme events? ....

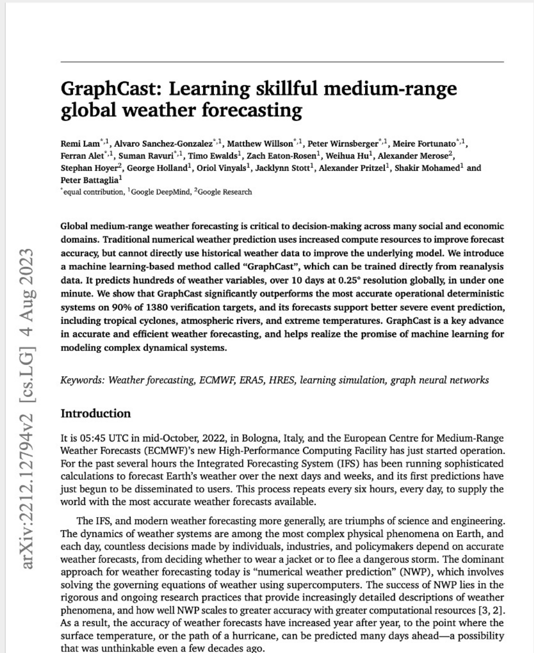
The deep learning revolution in weather forecasting

Slides adapted from David Landry, INRIA Paris (formerly of Environment Canada)

- Starting in 2023, several tech-industry research teams used deep neural networks to perform medium-range weather forecasts (e.g., NVIDIA, Huawei, Google DeepMind)
- The best ML-models' skill became ~equivalent to the leading Numerical Weather Prediction (NWP) models







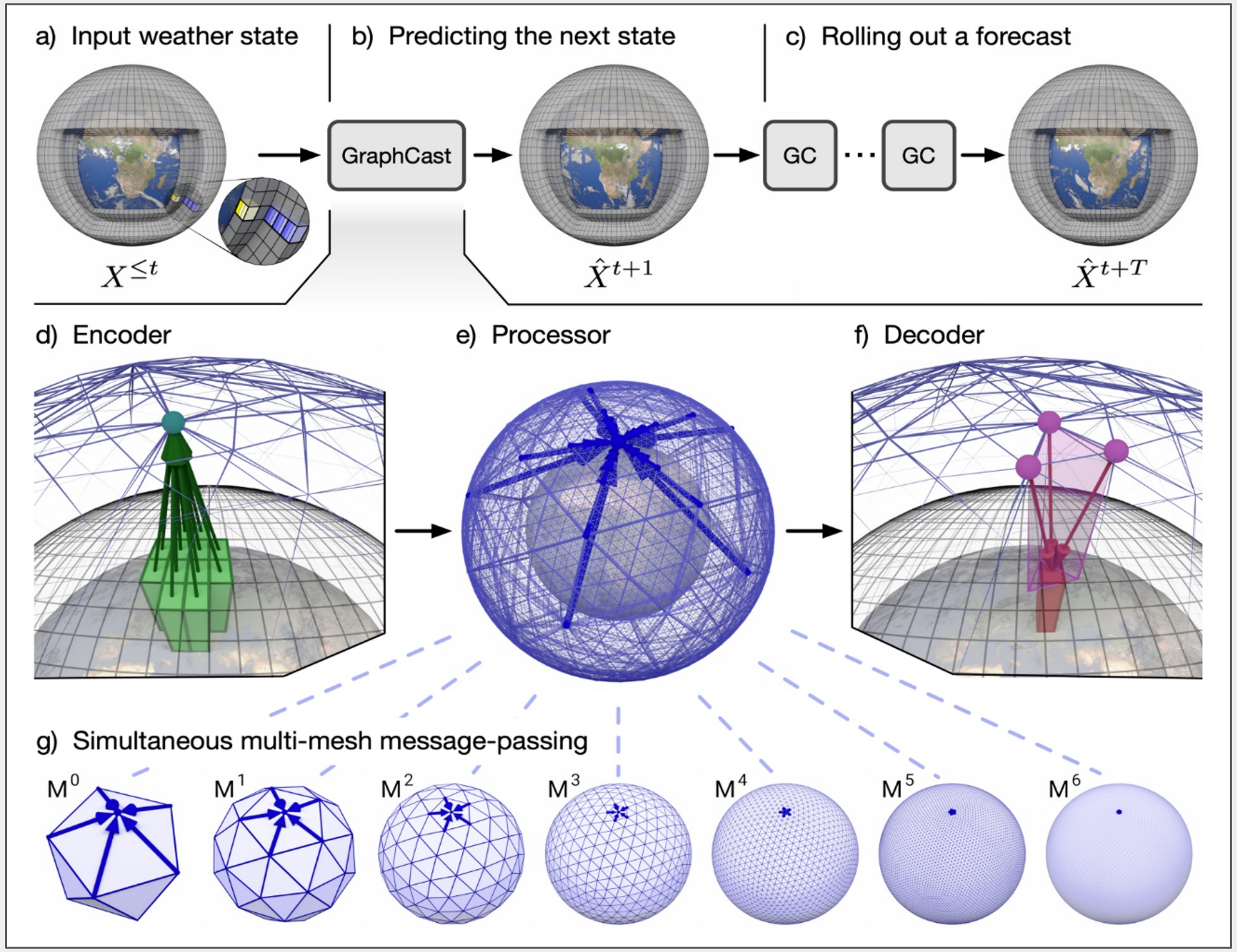
How it works

Models: Deep Neural Networks, e.g.:

- Fourier Neural Operators
- Graph Neural Networks
- Collection of models for different time scales

Training:

- Supervised learning
- Optimized for average performance
- Deterministic forecasts (not probabilistic)



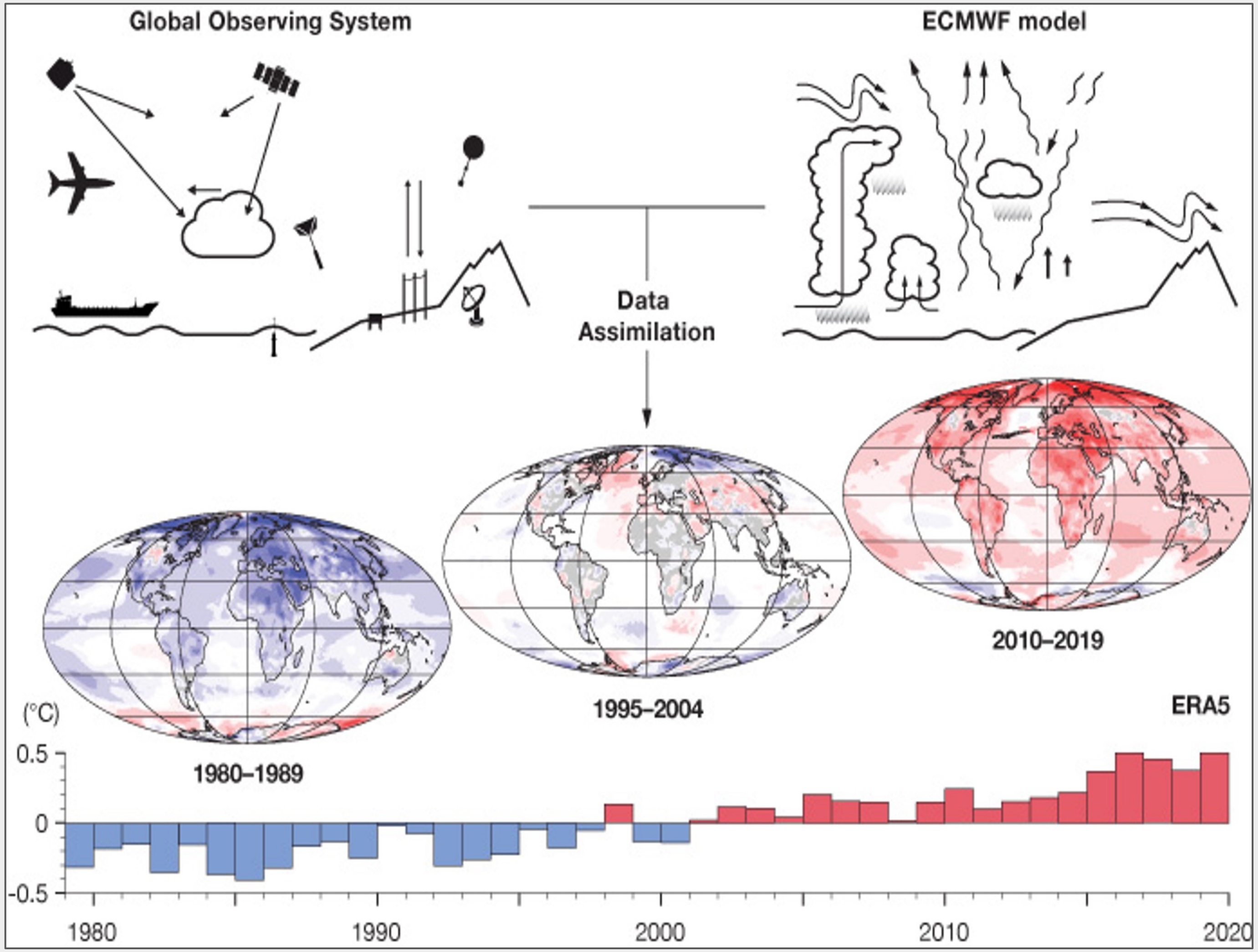
GraphCast architecture

doi:10.48550/arXiv.2212.12794

How it works

Training data: ERA5 reanalysis dataset

- Reanalysis: incorporating observations into a physically consistent system to create a best guess of the true state of the system



doi:10.1002/qj.3803
https://www.ecmwf.int/en/about/media-centre/focus/2023/fact-sheet-reanalysis

How it works

Model	Training	Inference
Pangu-Weather Huawei	16 days on 192 V100 GPUs	24 hours forecast 1xV100 1.4 seconds on
GraphCast Google DeepMind	21 days on 32 Cloud TPU V4 devices (~128 GPUs)	10 days forecast 1 Cloud TPU V4 < 60 seconds
NWP (ECMWF HRES) *Higher resolution	-	10 days forecast 1/3/6 hours 11 664 cores HPC cluster

ML weather models: Remaining challenges

Extreme event detection and forecasting

- ML “foundation models” were trained to predict averages
- Reanalysis data doesn’t capture some extreme events

Semi/Unsupervised learning

- Reduce reliance on Reanalysis data

Probabilistic forecasts

Will this scale from weather to longer (climate) time-scales?

Progress and remaining challenges

Addressing data limitations

- Climate models’ time/space scales are too course for many adaptation/mitigation tasks
- Class imbalance: e.g., extreme events are rare by definition!
- Limited *labeled* data

→ Unsupervised learning can help for all of the above!

- Data is limited along the time dimension
  - Can we substitute data diversity and granularity over space?

Non-stationarity

- Climate *change* means we cannot assume i.i.d. data!
- ML models need to adapt over time, and space

Interpretability

- Evaluation of generative models is an active research area of core ML

Long-term goals

Cascading Hazards

- Goal: move beyond individual weather extremes, to how they couple
- With massive wildfires everywhere, there is extreme urgency!

Climate Justice

- Our research should always help increase climate equity
- Ultimately, we should strive for approaches to help UNDO the legacy of climate IN-justice

Climate and Machine Learning Boulder (CLIMB)

Boulder NSF Climate Change AI futurearth Research. Innovation. Sustainability.

Thank you!

And many thanks to:

Arindam Banerjee, *University of Illinois Urbana-Champaign*
Nicolò Cesa-Bianchi, *Università degli Studi di Milano*
Tommaso Cesari, *Toulouse School of Economics*
Guillaume Charpiat, *INRIA Saclay*
Cécile Coléou, *Météo-France & CNRS*
Michael Dechartre, *Irstea, Université Grenoble Alpes*
Nicolas Eckert, *Irstea, Université Grenoble Alpes*
Brandon Finley, *University of Lausanne*
Sophie Giffard-Roisin, *IRD Grenoble*
Brian Groenke, *Alfred Wegener Institute, Potsdam*
Nidhin Harilal, *University of Colorado Boulder*
Tommi Jaakkola, *MIT*
Anna Karas, *Météo-France & CNRS*
Fatima Karbou, *Météo-France & CNRS*
Balázs Kégl, *Huawei Research & CNRS*
David Landry, *INRIA Paris*
Luke Madaus, *Jupiter Intelligence*
Scott McQuade, *Amazon*
Ravi S. Nanjundiah, *Indian Institute of Tropical Meteorology*
Moumita Saha, *Philips Research India*
Gavin A. Schmidt, *NASA Senior Advisor on Climate*
Saumya Sinha, *National Renewable Energy Lab*
Cheng Tang, *Amazon*

METEO FRANCE NASA

AI Research for Climate Change and Environmental Sustainability (ARCHES)

Inria

Choose France™

Environmental Data Science Innovation & Inclusion Lab

A national accelerator linking data, discovery, & decisions

ESIIL

NSF's newest data synthesis center,
hosted by the University of Colorado Boulder & CIRES,
with key partners CyVerse & the University of Oslo