Comments for round-table discussion

Traditional (physics-based) weather and climate modeling uses HPC

→ AI actually REDUCES compute in these domains!

Once trained, AI is significantly faster at prediction than physical models Likely the case for ML emulators of physical models in your domain!

We develop and train targeted models that are best for the applications

–> These are much smaller than huge foundation models e.g., from OpenAI, that are meant to be very general

You can do that in your domain as well!



Al Research for Climate Change and Environmental Sustainability

Claire Monteleoni

INRIA Paris

Inría Choose France[™]



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

"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: using Machine Learning to address 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
- 2024 13th Conference on Climate Informatics, Turing Institute, London
- 2025 14th Conference on Climate Informatics, April, Rio de Janeiro



Al Research for Climate Change and Environmental Sustainability



Approach: Exploit all available data

Simulated data generated by physics-based models

- Numerical Weather Prediction (NWP) models
- General Circulation Models (GCM)
- Regional Climate Models (RCM)

Reanalysis data

Gridded data products from <u>data assimilation</u>: applies physical laws to observations

Observation data

- □ Satellite remote sensing data
- In-situ data



Al Methods

Semi-supervised, unsupervised, self-supervised learning

- New methods for downscaling (super-resolution), interpolation of geospatial data
- New pretext tasks for self-supervised learning, e.g., STINT [Harilal et al., 2024]
- Regularization via multi-tasking over variables, lead-times

Generative AI

- □ VAE, Normalizing Flows
- Diffusion and flow-based training
- Develop new generative downscaling methods, e.g., [Groenke et al., 2020]

Learning under non-stationarity

□ Learn level of non-stationarity over time and space

ADAPTATION AI for Extreme Weather and Cascading Hazards



Hurricane track prediction

Forecasting Indian Summer Monsoon precipitation extremes

Avalanche detection

Generative AI for weather forecasting

[Giffard-Roisin et al., Frontiers 2020]



MITIGATION

Reducing carbon emissions

Accelerate green energy transition

- Al-driven forecasting of solar, wind
- AI to downscale solar and wind data

Reduce compute for weather and climate modeling

• Once trained, AI is significantly faster at prediction than physical models

IMPACTS Al for Understanding and Predicting Climate Change



Use AI to learn relations between IPCC simulations and observations

Robustify climate model ensemble forecasts

- Projecting long-term sea-level rise
- Projecting long-term carbon emissions

Al for downscaling spatiotemporal data

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

- Climate change adaptation
- Climate change mitigation
- Projecting long-term impacts

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



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

Revolution in AI for weather forecasting



Since 2022, a variety of AI models have shown weather forecasting performance comparable or **BETTER** than numerical weather prediction (NWP).

These deep learning (DL) models are trained on <u>reanalysis</u> data (ERA5) to predict the next weather state given the current state

Model predictions are then « rolled-out » to forecast 7-10 days in the future

What is reanalysis data?

Image credit: European Center for Medium Range Weather Forecasting (ECMWF) website

ArchesWeatherGen



Geopotential

Temperature

Humidity

Wind Speed

[Couairon, Singh, Charantonis, Lessig, Monteleoni, 2024]

Score cards on Google's Weather Bench 2

Generative AI for weather forecasting



[ArchesWeatherGen, Couairon, Singh, Charantonis, Lessig, Monteleoni, 2024]

Al for Climate Data Equity

- Train models in high-data regions and apply them in low-data regions
 - Train and validate them in high-data regions
 - Fine-tune them using the limited data in the low-data regions and use them to generate more data.
- Contribution to climate data equity
 - Local scales (e.g. legacy of environmental injustice in USA)
 - 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

Are Black Americans Underserved by the NWS Radar Network?

40-50%

30-40%

"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]



The farther a storm is from a radar site, the less information we can get about it due to the beam height rising farther off the ground, and the beam width expanding leading to lower resolution.

High resolution radar data near the ground can be critical in many situations such as when severe thunderstorms and tornadoes threaten.

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

Black Population Share

10-20%

20-30%

0-10%



70-80%

80-90%

90-100%

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)





futurearth

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







AI Research for Climate Change and Environmental Sustainability (ARCHES)



ENVIRONMENTAL DATA SCIENCE

An interdisciplinary, open access journal dedicated to the potential of artificial intelligence and data science to enhance our understanding of the environment, and to address climate change.

Data and methodological scope: Data Science broadly defined, including:

Machine Learning; Artificial Intelligence; Statistics; Data Mining; Computer Vision; Econometrics

Environmental scope, includes:

Water cycle, atmospheric science (including air quality, climatology, meteorology, atmospheric chemistry & physics, paleoclimatology)

Climate change (including carbon cycle, transportation, energy, and policy)

Sustainability and renewable energy (the interaction between human processes and ecosystems, including resource management, transportation, land use, agriculture and food)

Biosphere (including ecology, hydrology, oceanography, glaciology, soil science)

Societal impacts (including forecasting, mitigation, and adaptation, for environmental extremes and hazards) Environmental policy and economics

www.cambridge.org/eds







Environmental Data Science Innovation & Inclusion Lab

A national accelerator linking data, discovery, & decisions

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

