Hybrid Physics-AI Forecast Models

Sergey Frolov, Matt Newman, Colin Zarzycki

In collaboration with:
Co-Authors

Presented at:

Hybrid model:
A combination of ML and traditional physics-driven models

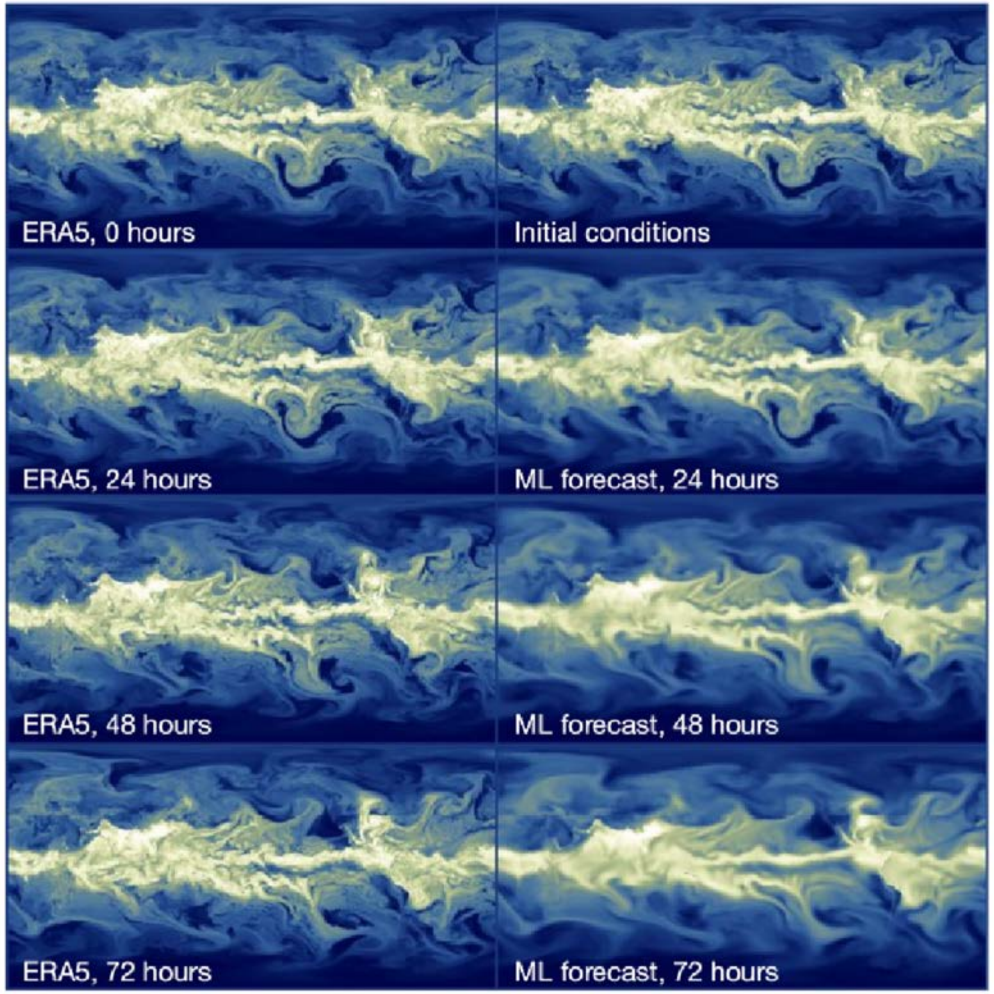

Figure 5: An example multi-step rollout of the ML forecast vs reanalysis data from ERA5. Beginning with the ERA5 initial conditions at 0 hours, the ML system steps forward autoregressively in 6-hour steps. While the model evolves 78 separate physical channels, we show only Q850, the specific humidity on the 850 hPa pressure level. The output of the ML forecast generally tracks the large-scale flows seen in ERA5, although the predictions do become smoother over time. Additional media, including videos, can be found at `https://rkeisler.github.io/graph_weather`

arXiv:2202.07575v1

- [right] Many efforts are underway to approximate the full non-linear model with a ML emulator.
- However, what are the best practices to use ML and traditional models in tandem to achieve best results?

Topics

- Integrating dynamical models with AI-based physics parameterization.
- Learning systematic model errors by combining model forecasts and observations.
- Using ML for climate prediction and climate event attribution, and to identify forecasts of opportunity.
- Limitations of the training data.

Talks

- Tapio Schneider, Caltech/NASA JPL
  - **AI-aided hybrid parameterizations**
- William Crawford, NRL
  - **Using analysis corrections to represent systematic and stochastic model error**
- John Albers, University of Colorado
  - **ML for climate prediction/attribution: Use and best practices**

Questions: plenary

- What are the current gaps in our knowledge of these two approaches (physics vs. AI models) compared to the hybrid approach? Does hybrid mean a mix of the two approaches within a single model, or can it also be useful to have hybrid ensembles of purely physical and AI models?

- What opportunities for stochastic modeling are opened up with the introduction of ML and hybrid models?

- Can we establish "best practices" for applying ML to climate problems?
  - What are the limitations of extrapolating from the training dataset to future forecasts? What are the community needs for open training datasets? [Where did the Data Science WG leave this effort?]
  - How do we test techniques under "perfect" conditions; e.g., should we apply ML to problems where the "answer" is already known?
  - Can we identify the conditions where a hybrid physics/AI model is more useful than either a pure physical or pure AI approach? If so, what are they, and if not, how might we do this?

End

Questions: POS

- What are the needs and opportunities for developing open datasets that can facilitate development of hybrid / ML models.
- Balance between using observational and model data to support science driven by ML-driven forecasts.
- How to integrate process studies in the development of ML / hybrid models.

Questions: PSMI

- How can hybrid ML approaches inform targets for observational data requirements?
- What is the state of interpretability with hybrid ML approaches and how can they improve our understanding of physical processes within the climate system?
- Where are challenges/gaps and associated best practices in physically-constrained optimization/loss?

Questions: PPAI

Can we make recommendations about some "best practices" for applying ML to climate prediction, and what should they cover?

- Establish null hypothesis, tests of overfitting and of technique (ie, does technique do what we expect it to do)
- How is climate change (non-stationary statistics) identified?
- Is the goal to create alternative forecasts or better forecasts?
- What does ensemble forecasting look like? How do we/should we identify forecasts of opportunity?
- Forecast models for weather/climate, or skip straight to applications?