Overview of AI Weather Predictions

How Does it Work and What are the Implications to Power System Applications?

Dr. Sue Ellen Haupt
Senior Scientist
Research Applications Lab, NCAR

Energy Systems Integration Group

June 24, 2025

Forecasting & Markets Workshop

ECMWF - 2018

Geosci. Model Dev., 11, 3999–4009, 2018
https://doi.org/10.5194/gmd-11-3999-2018

Geoscientific Model Development
Open Access
EGU

Challenges and design choices for global weather and climate models based on machine learning

Peter D. Dueben and Peter Bauer

European Centre for Medium-range Weather Forecasts, Shinfield Rd, Reading, RG2 9AX, UK

Correspondence: Peter D. Dueben (peter.dueben@ecmwf.int)

Received: 16 June 2018 – Discussion started: 28 June 2018
Revised: 30 August 2018 – Accepted: 12 September 2018 – Published: 1 October 2018

Abstract. Can models that are based on deep learning and trained on atmospheric data compete with weather and climate models that are based on physical principles and the basic equations of motion? This question has been asked often recently due to the boom in deep-learning techniques. The question is valid given the huge amount of data that are available, the computational efficiency of deep-learning techniques and the limitations of today's weather and climate models in particular with respect to resolution and complexity.

In this paper, the question will be discussed in the context of global weather forecasts. A toy model for global weather predictions will be presented and used to identify challenges and fundamental design choices for a forecast system based on neural networks.

1 Introduction

In recent years, artificial intelligence and machine learning have become very important for hardware development in high-performance computing (HPC) and have attracted a large amount of public interest. Neural networks (NNs) are tools from machine learning that are used successfully within many applications such as computer vision, speech recognition and data filtering. If a sufficient amount of data are available, NNs can be trained to describe the evolution of non-linear processes. Due to the fundamentally application-unaware character, no complete understanding of the underlying process is necessary. Very complex NNs can be trained that use more than a billion trainable parameters and millions of datasets for training on HPC architecture; see, for example, Le (2013).

On the other hand, numerical weather forecasts are computationally expensive and forecast quality reduces significantly already after a couple of days even in the best models available. Most processes in the Earth system are described by non-linear differential equations with non-linear interactions between Earth system components. Due to the complexity and size of the Earth system and the limited capacity of today's supercomputers, it is necessary to make approximations when weather prediction models are formulated and resolution is truncated in space and time. The use of limited resolution makes it necessary to parameterise processes that are not resolved explicitly within model simulations. To optimise parameterisation schemes a large number of parameters has to be tuned towards optimal model performance, and the traceability of physical laws of the underlying process as well as the physical interpretation for each parameter is often lost during this exercise. Furthermore, to perform weather predictions, a huge amount of data need to be processed and assimilated to create initial conditions. This is a process that will again cause significant errors and uncertainties. Only a rather small fraction of all observations can be assimilated into state-of-the-art weather prediction models due to the large computational cost and simplified assumptions required such as vanishing error correlation.

NNs have been used to post-process data from weather forecast models to optimise predictions; see, for example, Krasnopolsky and Lin (2012) or Rasp and Lerch (2018). NNs have also been used for radiation parameterisation in operational forecasts at ECMWF in the past (Chevallier et al., 1998, 2000; Krasnopolsky et al., 2005) as well as for the parameterisation of ocean physics (Krasnopolsky et al., 2002; Tolman et al., 2005) and convection (Krasnopolsky et al., 2013). Recently, the representation of atmospheric sub-grid

Published by Copernicus Publications on behalf of the European Geosciences Union.

It is possible to make global weather forecasts with a toy NN model that are better than persistence and competitive with T21 Atmospheric models of similar complexity for short lead times

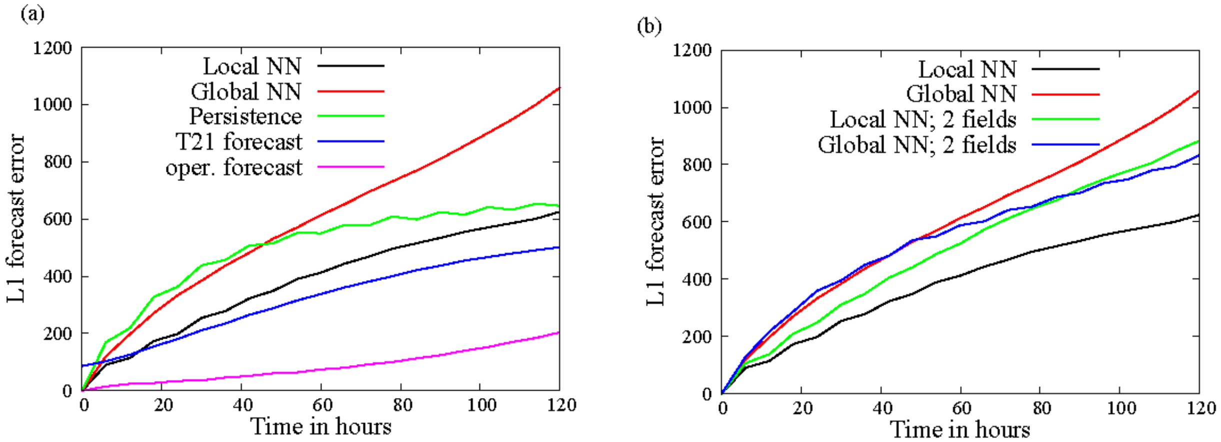


Figure 3. (a) Globally integrated absolute forecast error for the best local network (9×9 stencil), the global network, a persistence forecast, an IFS forecast at TL21 resolution and the operational weather forecast of ECMWF. The persistence forecast shows a 12-hourly fluctuation since Z500 has a weak 12-hourly cycle in the tropics due to atmospheric tides. **(b)** The same globally integrated absolute forecast error for the best local and global network as in **(a)** plus the best results for local and global networks that use 2mT as additional prognostic field.

Univ. Wash + MS - 2021

AGU ADVANCING EARTH AND SPACE SCIENCE

JAMES Journal of Advances in Modeling Earth Systems

RESEARCH ARTICLE
10.1029/2021MS002502

Sub-Seasonal Forecasting With a Large Ensemble of Deep-Learning Weather Prediction Models

Jonathan A. Weyn[1,3], Dale R. Durran[1], Rich Caruana[2], and Nathaniel Cresswell-Clay[1]

[1]Department of Atmospheric Sciences, University of Washington, Seattle, WA, USA, [2]Microsoft Research, Redmond, WA, USA, [3]Microsoft, Redmond, WA, USA

Key Points:
- An ensemble forecast system is developed using convolution neural networks (CNNs) to generate data-driven global forecasts
- Only 3 s are required to compute a large 320-member ensemble of skillful 6-week sub-seasonal predictions
- Shorter lead time forecasts also show skill, including a single deterministic 4-day forecast for Hurricane Irma

Correspondence to:
J. A. Weyn,
jweyn@uw.edu

Citation:
Weyn, J. A., Durran, D. R., Caruana, R., & Cresswell-Clay, N. (2021). Sub-seasonal forecasting with a large ensemble of deep-learning weather prediction models. *Journal of Advances in Modeling Earth Systems*, *13*, e2021MS002502. https://doi.org/10.1029/2021MS002502

Received 9 FEB 2021
Accepted 19 JUN 2021

Abstract We present an ensemble prediction system using a Deep Learning Weather Prediction (DLWP) model that recursively predicts six key atmospheric variables with six-hour time resolution. This computationally efficient model uses convolutional neural networks (CNNs) on a cubed sphere grid to produce global forecasts. The trained model requires just three minutes on a single GPU to produce a 320-member set of six-week forecasts at 1.4° resolution. Ensemble spread is primarily produced by randomizing the CNN training process to create a set of 32 DLWP models with slightly different learned weights. Although our DLWP model does not forecast precipitation, it does forecast total column water vapor and gives a reasonable 4.5-day deterministic forecast of Hurricane Irma. In addition to simulating mid-latitude weather systems, it spontaneously generates tropical cyclones in a one-year free-running simulation. Averaged globally and over a two-year test set, the ensemble mean RMSE retains skill relative to climatology beyond two-weeks, with anomaly correlation coefficients remaining above 0.6 through six days. Our primary application is to subseasonal-to-seasonal (S2S) forecasting at lead times from two to six weeks. Current forecast systems have low skill in predicting one- or 2-week-average weather patterns at S2S time scales. The continuous ranked probability score (CRPS) and the ranked probability skill score (RPSS) show that the DLWP ensemble is only modestly inferior in performance to the European Center for Medium Range Weather Forecasts (ECMWF) S2S ensemble over land at lead times of 4 and 5–6 weeks. At shorter lead times, the ECMWF ensemble performs better than DLWP.

Plain Language Summary The world's leading weather forecasting institutions currently rely on computationally expensive weather models running on massive supercomputers. In order to have predictive skill for forecasts two to six weeks in the future, large ensembles of many nearly identical runs of these models are required, but the computational resources needed for these ensembles scales with the number of forecasts run. Since the resources needed rapidly approaches modern-day computing limits, we explore the possibility of using computationally cheap weather models based on machine learning algorithms which learn to reproduce the evolution of weather. Our machine-learning model is capable of running 320 forecasts in three minutes on a single workstation, while the state-of-the-art model from the European Center for Medium-Range Weather Forecasts (ECMWF) utilizes supercomputers to run 50 forecasts. Our ensemble weather model produces realistic forecasts of weather events such as Hurricane Irma in 2017 and is even capable of nearly matching the performance of the ECMWF ensemble for forecasts of temperature four to six weeks in the future.

1. Introduction

Weather forecasting relies heavily on data assimilation to estimate the current state of the atmosphere and on numerical weather prediction (NWP) to approximate its subsequent evolution. The skill of such deterministic weather forecasts is typically limited to about two weeks by the chaotic growth of small initial errors and inaccuracies in our approximate models of the atmosphere. On much longer, multi-month time scales, the coupling of the atmosphere with slowly evolving ocean-land forcing allows skillful seasonal forecasts of monthly or seasonally averaged conditions. Between these two extremes, the production of skillful one- or two-week averaged forecasts at lead times ranging roughly between two weeks and two months (the subseasonal-to-seasonal or S2S time frame) has proven particularly challenging; yet there are many societal sectors that would greatly benefit from improved S2S forecasts (White et al., 2017). Several major operational centers have developed NWP-based ensemble systems focused on improving S2S forecasting (Vitart et al., 2017).

- Built deep-learning-based convolutional neural network ensemble system for S2S forecasting.
- Requires 3 min to produce a 320-member 6-wk ensemble forecast
- Similar scores (CRPS and RPSS) for 4-wk fx/ and 5-6-wk fx/ as ECMWF S2S ensembles.

AGU ADVANCING EARTH AND SPACE SCIENCE Journal of Advances in Modeling Earth Systems 10.1029/2021MS002502

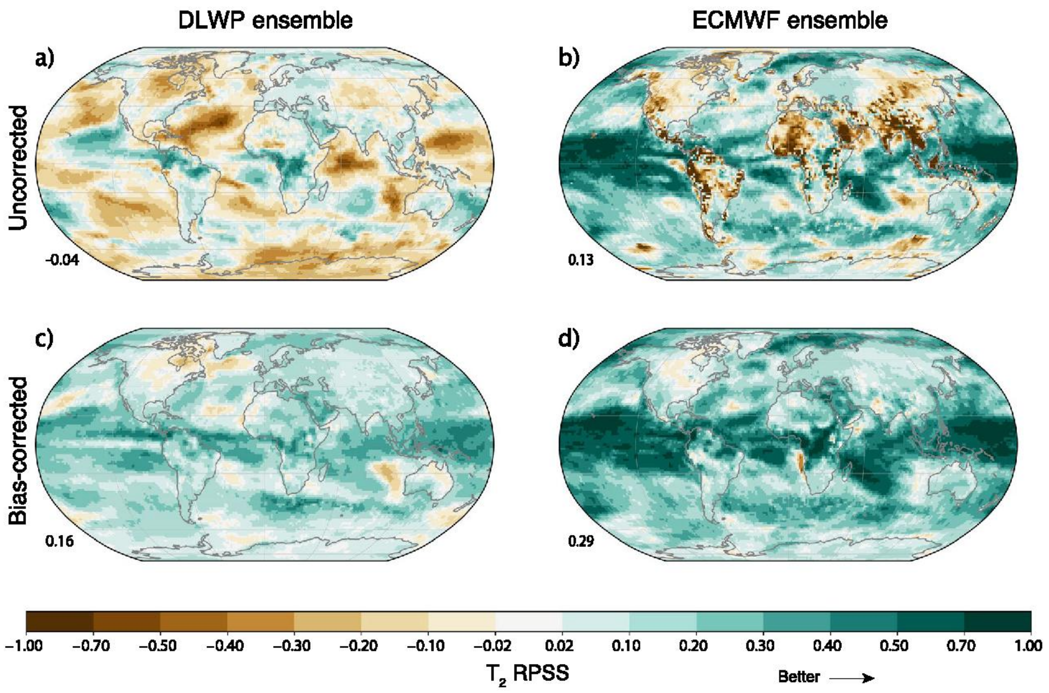


Figure 13. Annual average RPSS skill maps for T_2 at weeks 5–6. Without bias correction: (a) DLWP ensemble, (b) ECMWF ensemble; with bias correction: (c) DLWP ensemble, (d) ECMWF ensemble. The weighted global mean is noted at the lower left in each panel.

Subsequent Proliferation of Papers

GraphCast: Learning skillful medium-range global weather forecasting

Google – Aug. 2023

Remi Lam[*,1], Alvaro Sanchez-Gonzalez[*,1], Matthew Willson[*,1], Peter Wirnsberger[*,1], Meire Fortunato[*,1], Ferran Alet[*,1], Suman Ravuri[*,1], Timo Ewalds[1], Zach Eaton-Rosen[1], Weihua Hu[1], Alexander Merose[2], Stephan Hoyer[2], George Holland[1], Oriol Vinyals[1], Jacklynn Stott[1], Alexander Pritzel[1], Shakir Mohamed[1] and Peter Battaglia[1]

[*]equal contribution, [1]Google DeepMind, [2]Google Research

Global medium-range weather forecasting is critical to decision-making across many social and economic domains. Traditional numerical weather prediction uses increased compute resources to improve forecast accuracy, but cannot directly use historical weather data to improve the underlying model. We introduce a machine learning-based method called "GraphCast", which can be trained directly from reanalysis data. It predicts hundreds of weather variables, over 10 days at 0.25° resolution globally, in under one minute. We show that GraphCast significantly outperforms the most accurate operational deterministic systems on 90% of 1380 verification targets, and its forecasts support better severe event prediction, including tropical cyclones, atmospheric rivers, and extreme temperatures. GraphCast is a key advance in accurate and efficient weather forecasting, and helps realize the promise of machine learning for modeling complex dynamical systems.

Keywords: Weather forecasting, ECMWF, ERA5, HRES, learning simulation, graph neural networks

FourCastNet: A Global Data-driven High-resolution Weather Model using Adaptive Fourier Neural Operators

A Preprint **NVIDEA+ – Feb. 2022**

Jaideep Pathak, NVIDIA Corporation, Santa Clara, CA 95051
Shashank Subramanian, Lawrence Berkeley National Laboratory, Berkeley, CA 94720
Peter Harrington, Lawrence Berkeley National Laboratory, Berkeley, CA 94720
Sanjeev Raja, University of Michigan, Ann Arbor, MI 48109
Ashesh Chattopadhyay, Rice University, Houston, TX 77005
Morteza Mardani, NVIDIA Corporation, Santa Clara, CA 95051
Thorsten Kurth, NVIDIA Corporation, Santa Clara, CA 95051
David Hall, NVIDIA Corporation, Santa Clara, CA 95051
Zongyi Li, California Institute of Technology, Pasadena, CA 91125, NVIDIA Corporation, Santa Clara, CA 95051
Kamyar Azizzadenesheli, Purdue University, West Lafayette, IN 47907
Pedram Hassanzadeh, Rice University, Houston, TX 77005
Karthik Kashinath, NVIDIA Corporation, Santa Clara, CA 95051
Animashree Anandkumar, California Institute of Technology, Pasadena, CA 91125, NVIDIA Corporation, Santa Clara, CA 95051

February 24, 2022

TECHNICAL REPORT **Huawei – Nov. 2022**

Pangu-Weather: A 3D High-Resolution System for Fast and Accurate Global Weather Forecast

Kaifeng Bi, Lingxi Xie, Hengheng Zhang, Xin Chen, Xiaotao Gu, and Qi Tian, *Fellow, IEEE*

Abstract—In this paper, we present Pangu-Weather, a deep learning based system for fast and accurate global weather forecast. For this purpose, we establish a data-driven environment by downloading 43 years of hourly global weather data from the 5th generation of ECMWF reanalysis (ERA5) data and train a few deep neural networks with about 256 million parameters in total. The spatial resolution of forecast is 0.25° × 0.25°, comparable to the ECMWF Integrated Forecast Systems (IFS). More importantly, for the first time, an AI-based method outperforms state-of-the-art numerical weather prediction (NWP)

Prithvi WxC: Foundation Model for Weather and Climate

IBM/+ – Sept 2024

Johannes Schmude[1,†,‡], Sujit Roy[2,7,†,‡], Will Trojak[1], Johannes Jakubik[1], Daniel Salles Civitarese[1], Shraddha Singh[1], Julian Kuehnert[1], Kumar Ankur[2], Aman Gupta[3], Christopher E Phillips[2], Romeo Kienzler[1], Daniela Szwarcman[1], Vishal Gaur[2], Rajat Shinde[2], Rohit Lal[2], Arlindo Da Silva[6], Jorge Luis Guevara Diaz[1], Anne Jones[1], Simon Pfreundschuh[4], Amy Lin[2], Aditi Sheshadri[3], Udaysankar Nair[2], Valentine Anantharaj[5], Hendrik Hamann[1], Campbell Watson[1], Manil Maskey[7], Tsengdar J Lee[8], Juan Bernabe Moreno[1], Rahul Ramachandran[7]

[†]Equal Contribution,
[‡]Johannes.Schmude@ibm.com , Sujit.Roy@nasa.gov

Microsoft/UCLA – July2023

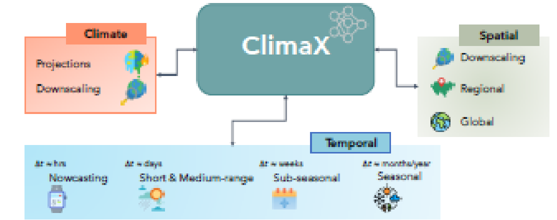
2023-7-12

ClimaX: A foundation model for weather and climate

Tung Nguyen[1], Johannes Brandstetter[2], Ashish Kapoor[3], Jayesh K. Gupta[*2], and Aditya Grover[*1]

[1]UCLA, [2]Microsoft, [3]Scaled Foundations

Most state-of-the-art approaches for weather and climate modeling are based on physics-informed numerical models of the atmosphere. These approaches aim to model the non-linear dynamics and complex interactions between multiple variables, which are challenging to approximate. Additionally, many such numerical models are computationally intensive, especially when modeling the atmospheric phenomenon at a fine-grained spatial and temporal resolution. Recent data-driven approaches based on machine learning instead aim to directly solve a downstream forecasting or projection task by learning a data-driven functional mapping using deep neural networks. However, these networks are trained using curated and homogeneous climate datasets for specific spatiotemporal tasks, and thus lack the generality of numerical models. We develop and demonstrate ClimaX, a flexible and generalizable deep learning model for weather and climate science that can be trained using heterogeneous datasets spanning different variables, spatio-temporal coverage, and physical groundings. ClimaX extends the Transformer architecture with novel encoding and aggregation blocks that allow ive use of available compute while maintaining general utility. ClimaX is pre-trained a self-supervised learning objective on climate datasets derived from CMIP6. The pre-ed ClimaX can then be fine-tuned to address a breadth of climate and weather tasks, ding those that involve atmospheric variables and spatio-temporal scales unseen during aining. Compared to existing data-driven baselines, we show that this generality in ClimaX s in superior performance on benchmarks for weather forecasting and climate projections, when pretrained at lower resolutions and compute budgets. Source code is available at ://github.com/microsoft/ClimaX.



1: ClimaX is built as a foundation model for any weather and climate modeling task. On the weather these tasks include standard forecasting tasks for various lead-time horizons at various resolutions, both y or regionally. On the climate front, making long term projections and obtaining downscaling results ower resolution model outputs are standard tasks.

Compare to ECMWF

European Centre for Medium Range Forecasting

How skilful are the latest ML-based weather forecasts?

First, the headline scores of the released ML-based models hold up to independent evaluation. When assessed with deterministic scores, such as root mean square error (RMSE) or anomaly correlation coefficient (ACC), Pangu-Weather is a legitimate rival for the IFS (see Figure 1 for example). This holds true not only when assessed against analyses, but also against observations, and when using the same initial condition as the IFS (as opposed to initialising from ERA5, which is done in the public papers).

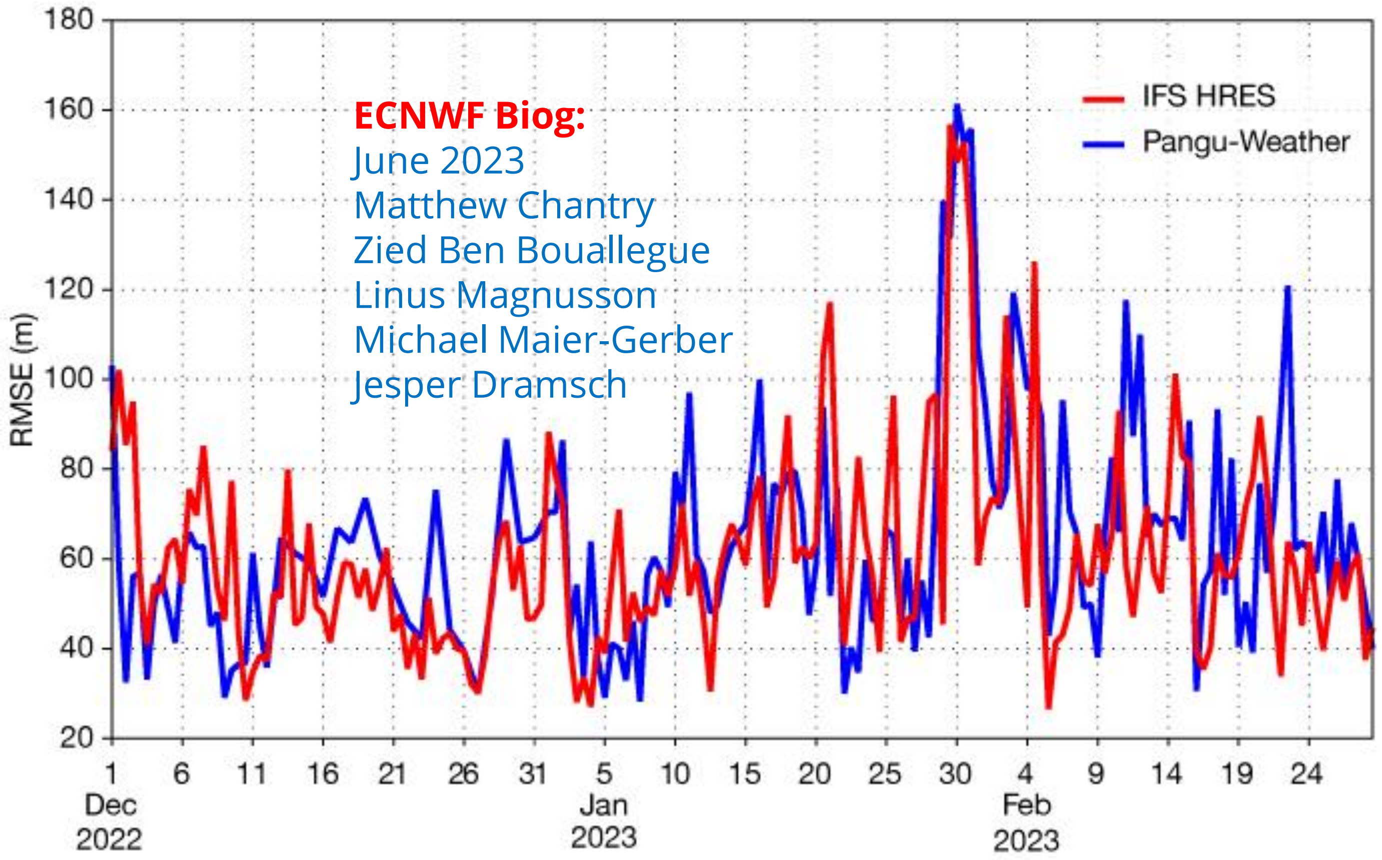


Figure 1: Root mean square error (RMSE) scores of 500 hPa geopotential height for IFS high-resolution forecasts (HRES) and Pangu-Weather over Europe for winter 2022/23 at day 6, measured against operational analysis. Pangu-Weather and the IFS produce comparably accurate forecasts and share a forecast "bust" near the end of January.

However, scores can be optimised, and ML models are trained to do exactly this. Pangu-Weather and FourCastNet were trained to minimise RMSE. Training towards this type of objective can smooth out predictions and it penalises forecasts of extremes. But of course, weather forecasts are at their most valuable for extreme events where lives are at stake.

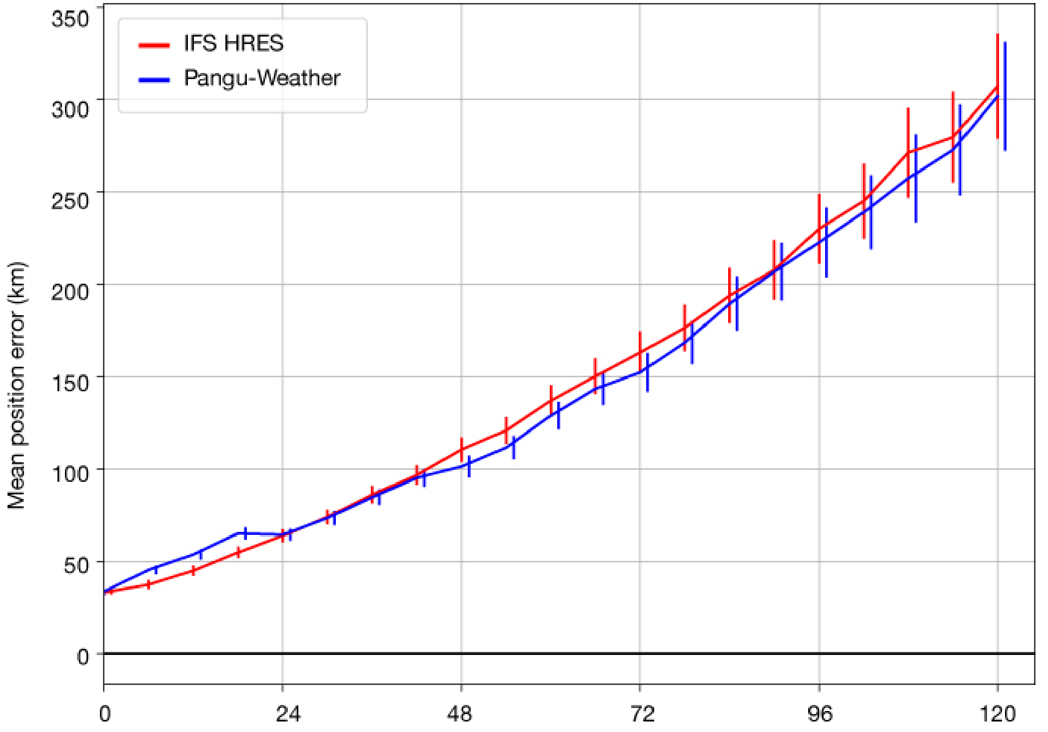


Figure 2: Average tropical cyclone track errors during 2018 for IFS high-resolution forecasts (HRES) and Pangu-Weather, measured against IBTrACS. The statistic is based on events having a tropical storm strength of at least 17m/s, and bars highlight the 95% confidence interval.

Does it Know the Dynamics?

Dynamical Tests of a Deep-Learning Weather Prediction Model

Sept. 2023

Gregory J. Hakim[1] and Sanjit Masanam[2]

[1]Department of Atmospheric Sciences, University of Washington, Seattle, WA
[2]Department of Physics, University of California at Santa Barbara, Santa Barbara, CA

Key Points:

- The Pangu-weather deep-learning weather prediction model exhibits physically realistic dynamical behavior
- Steady tropical heating produces a Matsuno–Gill response in the tropics, and planetary waves that radiate into the extratropics
- Localized initial conditions produce realistic hurricanes, extratropical cyclones, and adjustment to geostrophic balance

We conclude that the model encodes realistic physics in all experiments, and suggest it can be used as a tool for rapidly testing ideas before using expensive physics-based models.

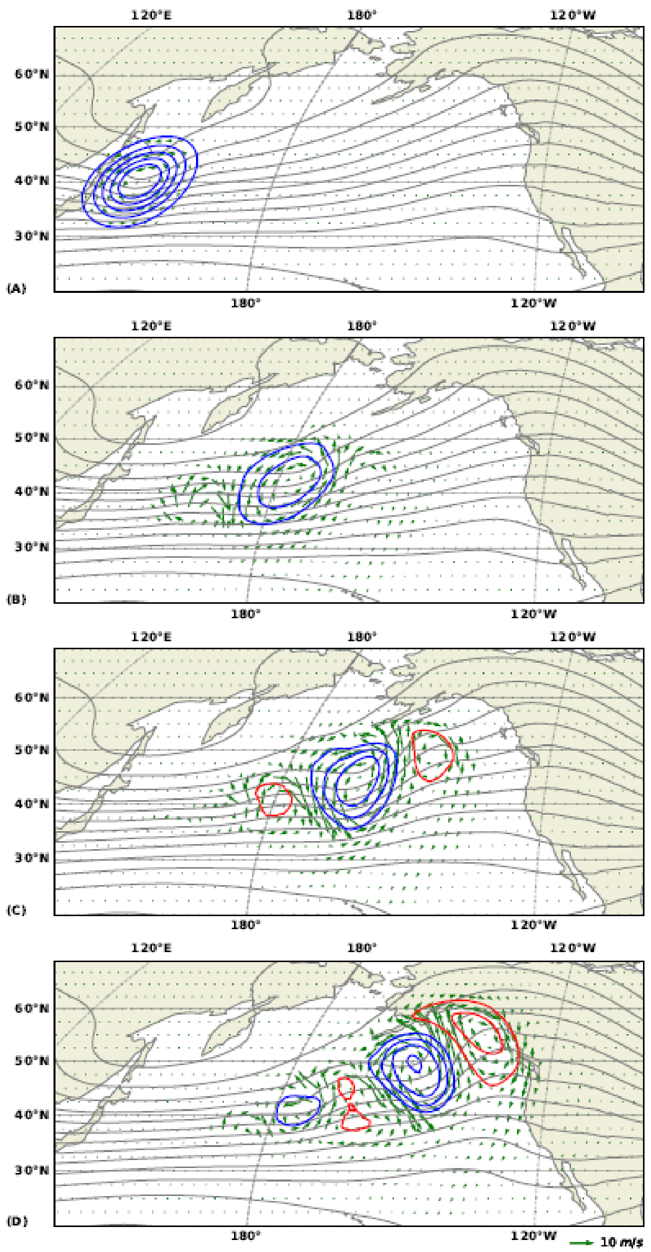


Figure 3. Solution at 500hPa for a localized disturbance on the DJF atmosphere. The full geopotential height is shown by gray lines every 60m, and anomalies from the DJF average by red (positive) and blue (negative) lines every 20m; the zero contour is suppressed. Green arrows show the anomalous vector wind. Solutions are shown at (A) 0 days (the specified initial condition); (B) 2 days; (C) 3 days; and (D) 4 days.

Can AIWP be Trained from Observations?

DATA DRIVEN WEATHER FORECASTS TRAINED AND INITIALISED DIRECTLY FROM OBSERVATIONS

A PREPRINT

Anthony McNally Christian Lessig Peter Lean Eulalie Boucher Mihai Alexe

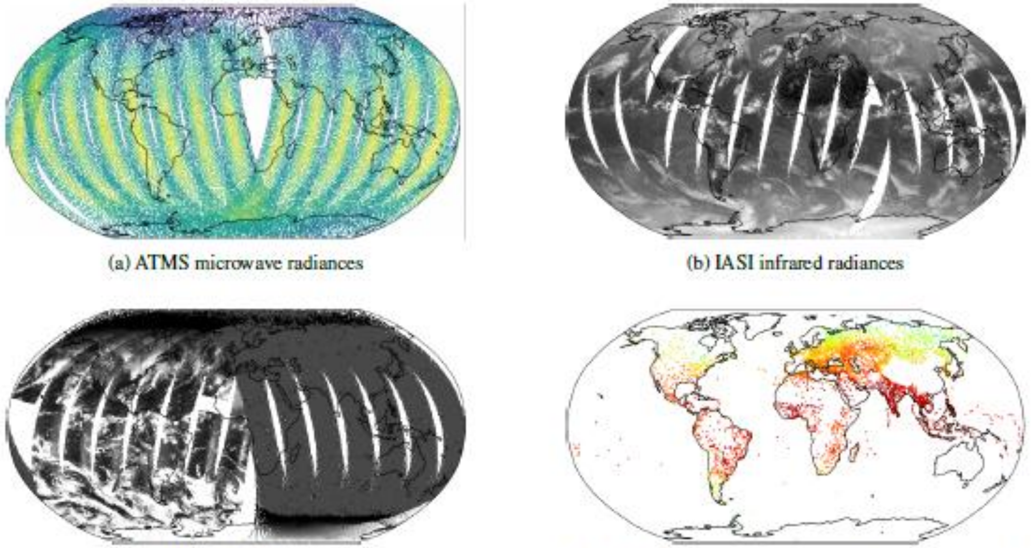
Ewan Pinnington Matthew Chantry Simon Lang Chris Burrows Marcin Chrust

Florian Pinault Ethel Villeneuve Niels Bormann Sean Healy

European Centre for Medium-Range Weather Forecasts (ECMWF)

ECMWF - July 2024

July 23, 2024



(a) ATMS microwave radiances

(b) IASI infrared radiances

(c) AVHRR visible reflectances

(d) SYNOP surface measurements of 2m temperature and 10m wind

Figure 1: Typical examples of the data coverage provided by ATMS microwave radiances (a), IASI infrared radiances (b), AVHRR visible reflectances (c) and SYNOP surface measurements of 2m temperature and 10m wind (d) within a 12-hour window.

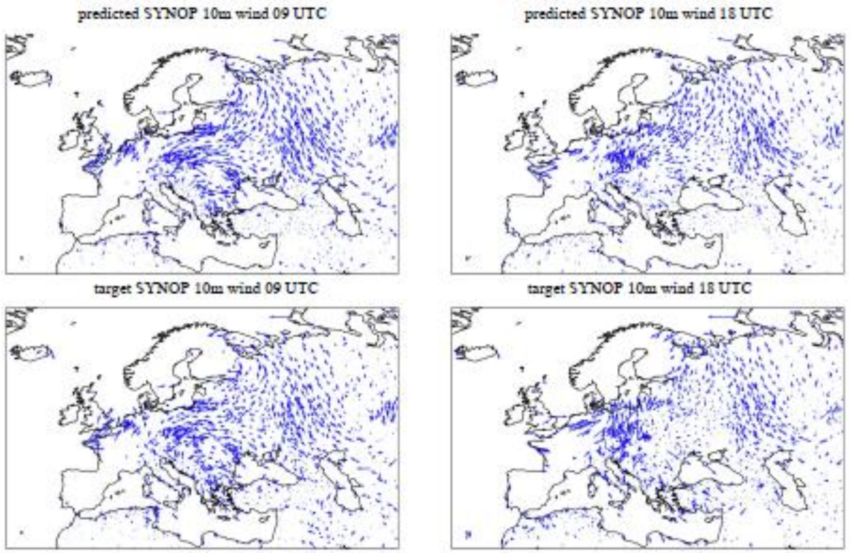


Figure 5: Example of 10m winds inferred from satellite brightness temperatures alone. Predicted SYNOP 10m wind (upper panels) and verifying target SYNOP 10m wind (lower panels) at 9UTC and 18UTC for a case on February 18th 2022.

TESTING THE LIMIT OF ATMOSPHERIC PREDICTABILITY WITH A MACHINE LEARNING WEATHER MODEL

A PREPRINT

P. Trent Vonich *[1,2] and Gregory J. Hakim †[1]
[1]Department of Atmospheric Sciences, University of Washington, Seattle, WA, USA
[2]Air Force Institute of Technology, Wright-Patterson AFB, OH, USA

April 30, 2025

ABSTRACT

Atmospheric predictability research has long held that the limit of skillful deterministic weather forecasts is about 14 days. We challenge this limit using GraphCast, a machine-learning weather model, by optimizing forecast initial conditions using gradient-based techniques for twice-daily forecasts spanning 2020. This approach yields an average error reduction of 86% at 10 days, with skill lasting beyond 30 days. Mean optimal initial-condition perturbations reveal large-scale, spatially coherent corrections to ERA5, primarily reflecting an intensification of the Hadley circulation. Forecasts using GraphCast-optimal initial conditions in the Pangu-Weather model achieve a 21% error reduction, peaking at 4 days, indicating that analysis corrections reflect a combination of both model bias and a reduction in analysis error. These results demonstrate that, given accurate initial conditions, skillful deterministic forecasts are consistently achievable far beyond two weeks, challenging long-standing assumptions about the limits of atmospheric predictability.

Newsletter

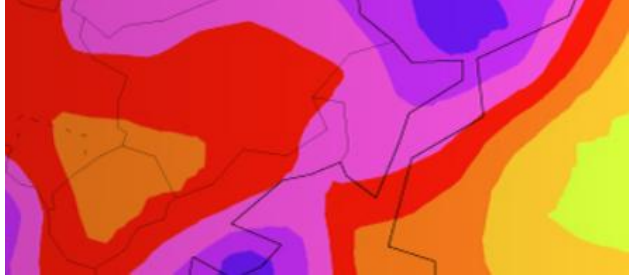
Number 178 - Winter 2024

Published in January 2024

NEWS

AIFS: a new ECMWF forecasting system

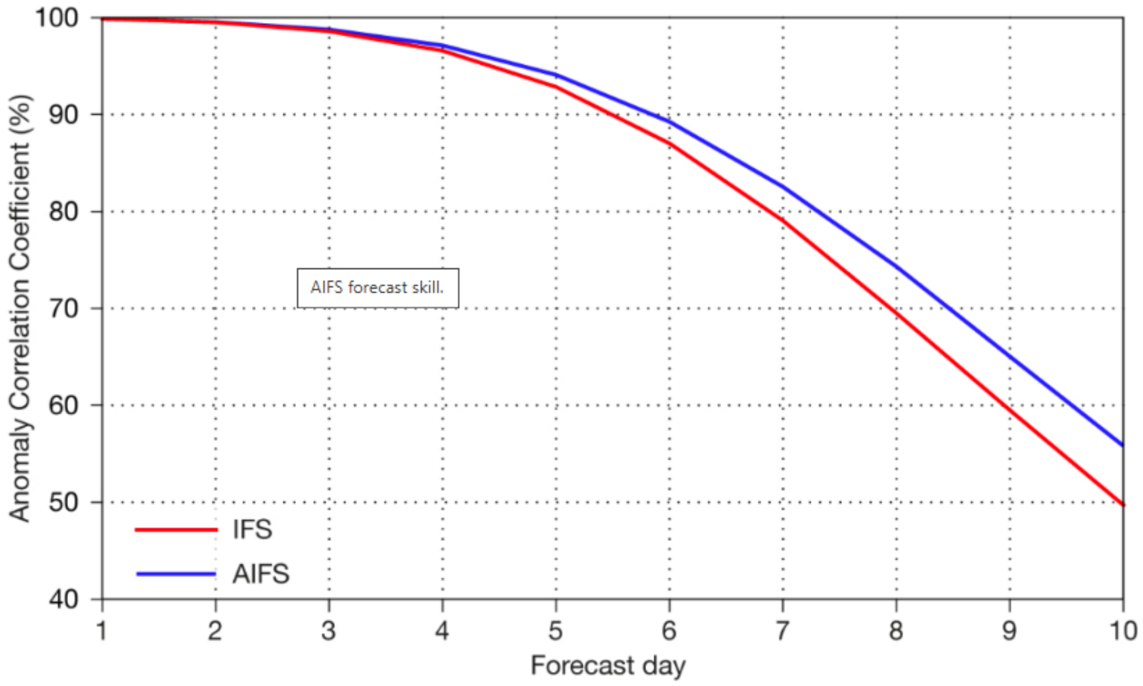
Simon Lang, Mihai Alexe, Matthew Chantry, Jesper Dramsch, Florian Pinault, Baudouin Raoult, Zied Ben Bouallègue, Mariana Clare, Christian Lessig, Linus Magnusson, Ana Prieto Nemesio



ECMWF has taken the Artificial Intelligence Forecasting System (AIFS) into operations today, 25 February 2025, to run side by side with its traditional physics-based Integrated Forecasting System (IFS) to advance numerical weather prediction.

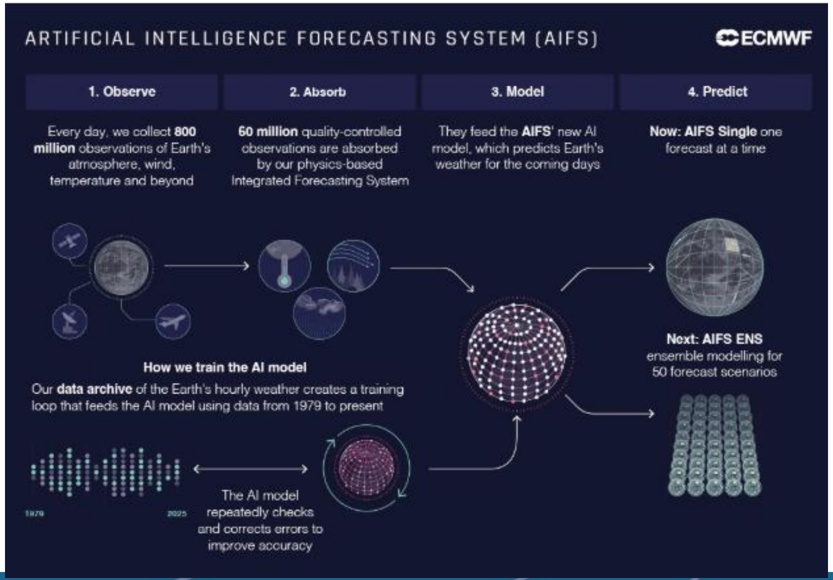
The AIFS outperforms state-of-the-art physics-based models for many measures, including tropical cyclone tracks, with gains of up to 20%.

This high-accuracy model complements the portfolio of our physics-based models by leveraging the opportunities made available by machine learning (ML) and artificial intelligence (AI). These include increased speed and a reduction of approximately 1,000 times in energy use for making a forecast.



AIFS forecast skill. We show the northern hemisphere Anomaly Correlation Coefficient (ACC) for geopotential height at 500 hPa of IFS forecasts (red, dashed) and AIFS forecasts (blue) for 2022. Higher values indicate better skill.

https://www.ecmwf.int/en/about/media-centre/news/2025/ecmwfs-ai-forecasts-become-operational



https://www.ecmwf.int/en/about/media-centre/news/2025/ecmwfs-ai-forecasts-become-operational

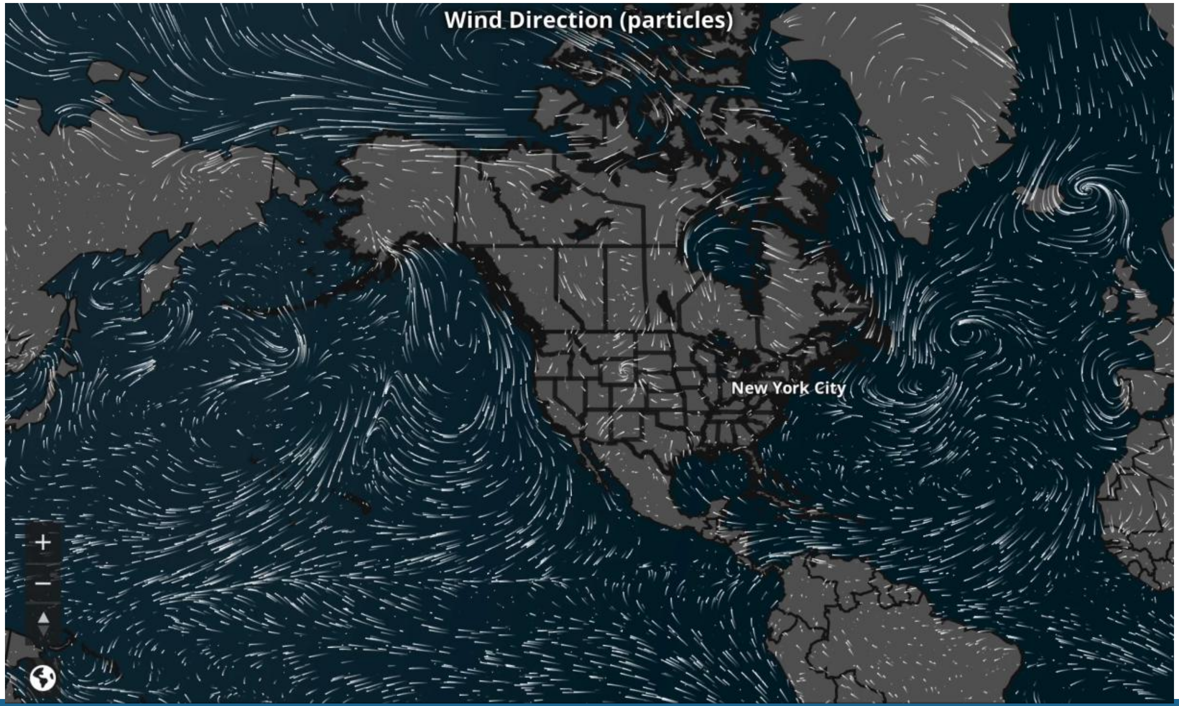
What is NOAA Doing?

An early look at NOAA's Project EAGLE to accelerate AI weather prediction advances for the United States

June 18, 2025

Authors: Sergey Frolov, Jun Wang, Isidora Jankov, Jacob Carley, Keven Blackman, Daryl Kleist, Travis Wilson, Linlin Cui, John Ten Hoeve and Maoyi Huang

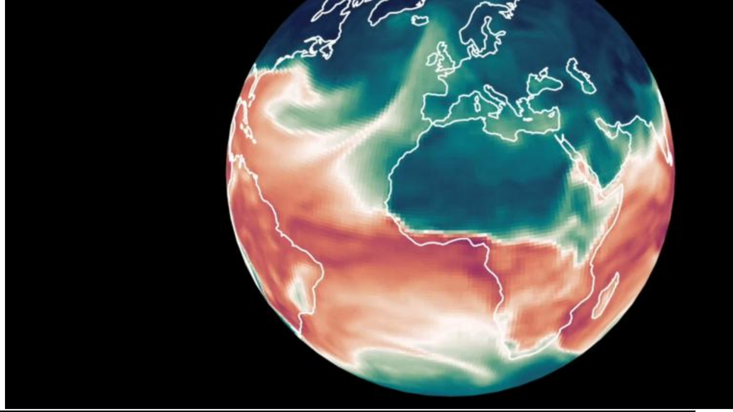
Today we are sharing some early progress on Project EAGLE (Experimental AI Global and Limited-area Ensemble forecast system), which is a joint effort between NOAA Research Laboratories and the Earth Prediction Innovation Center (EPIC) in the Office of Oceanic and Atmospheric Research (OAR), and the National Weather Service (NWS).



Today EAGLE includes two components: Global-EAGLE-Solo and Global-EAGLE-Ensemble. Both of them are based on Google DeepMind's GraphCast model (Lam et al., 2023) and are tuned by the NOAA Environmental Modeling Center (EMC) using NOAA data. The model runs on a 0.25-degree latitude-longitude grid (about 28 km) and 13 pressure levels. The model produces 16-day forecasts 2 times a day at 00Z and 12Z.

- **Global-EAGLE-Solo:** A demonstration environment for "deterministic" models that are initialized from a single GFS initial condition (IC). It is a complement to the existing NOAA GFS physics-based forecast. NOAA EMC re-trained the GraphCast model using Global Data Assimilation System (GDAS) data as inputs and training targets (Tabas et al., 2025).
- **Global-EAGLE-Ensemble:** A demonstration environment for ensemble forecasts, driven by the ICs of the operational GEFS. It is a complement to the existing NOAA GEFS physics-based ensemble forecast system. The weights for the Global-Eagle-Ensemble members are generated by fine-tuning the original GraphCast weights from DeepMind(c) with recent NOAA operational GDAS analysis. The resulting weights are effectively trained on the combination of European Centre for Medium-Range Weather Forecasts's (ECMWF) fifth-generation reanalysis (ERA5), ECMWF high-resolution (HRES), and NOAA GDAS analysis. Multiple checkpoints were saved to form 31 global ensemble members (Wang et al., 2025).

NSF NCAR Models



Community Research Earth Digital Intelligence Twin (CREDIT) *

NSF NCAR – Nov 2024

John S. Schreck[♠†‡], Yingkai Sha[♠†‡], William Chapman[♠†§], Dhamma Kimpara[†], Judith Berner[§], Seth McGinnis[†], Arnold Kazadi[†], Negin Sobhani[‡], Ben Kirk[‡], David John Gagne II[†‡]

Machine Integration and Learning for Earth Systems (MILES) Group[†]
Computational and Information Systems (CISL) Lab[‡]
Climate and Global Dynamics (CGD) Lab[§]

NSF National Center for Atmospheric Research
Boulder, Colorado, USA

Improving AI weather prediction models using global mass and energy conservation schemes *

NSF NCAR – Jan 2025

Yingkai Sha[†], John S. Schreck[†], William Chapman[‡], David John Gagne II[†]

Computational and Information Systems (CISL) Laboratory[†]
Climate and Global Dynamics (CGD) Laboratory[‡]

NSF National Center for Atmospheric Research
Boulder, Colorado, USA

CAMulator: Fast Emulation of the Community Atmosphere Model *

NSF NCAR – Apr 2025

William E. Chapman[‡], John S. Schreck[†], Yingkai Sha[†], David John Gagne II[†], Dhamma Kimpara[†], Laure Zanna[ς], Kirsten J. Mayer[‡], Judith Berner[‡]

Climate and Global Dynamics (CGD) Laboratory[‡]
Computational and Information Systems (CISL) Laboratory[†]
Courant Institute of Mathematical Sciences, New York University[ς]

NSF National Center for Atmospheric Research
Boulder, Colorado, USA

Investigating the contribution of terrain-following coordinates and conservation schemes in AI-driven precipitation forecasts *

NSF NCAR – Mar 2025

Yingkai Sha[†], John S. Schreck[†], William Chapman[‡], David John Gagne II[†]

Computational and Information Systems (CISL) Laboratory[†]
Climate and Global Dynamics (CGD) Laboratory[‡]

NSF National Center for Atmospheric Research
Boulder, Colorado, USA

NSF NCAR Framework: CREDIT

What is CREDIT?

An open foundational platform for developing and deploying AI weather and Earth system prediction models.

CREDIT enables users to build custom data and modeling pipelines to load data, train configurable AI forward models, and deploy them for real-time forecasting, hindcasting, or scenario projections.

CREDIT offers both scientifically validated model configurations and endless customization for any use case.

Datasets

- ERA5
- Community Atmosphere Model
- GFS
- CONUS 404

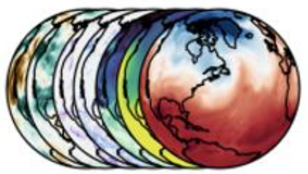
Models

- WXFormer
- NCAR FuXi
- Graph Recurrent Transformer

Physics

- Mass Conservation
- Energy Conservation
- Moisture Budget

Outputs



Thanks to David John Gagne

What about at finer Scales?

Geophysical Research Letters

Univ OK – Mar 2025

RESEARCH LETTER
10.1029/2024GL112383

WoFSCast: A Machine Learning Model for Predicting Thunderstorms at Watch-to-Warning Scales

Montgomery L. Flora[1,2] and Corey Potvin[2,3]

[1]Cooperative Institute for Severe and High-Impact Weather Research and Operations, University of Oklahoma, Norman, OK, USA, [2]NOAA/OAR/National Severe Storms Laboratory, Norman, OK, USA, [3]School of Meteorology, University of Oklahoma, Norman, OK, USA

Key Points:

- WoFSCast is an AI-numerical weather prediction system trained to predict rapidly evolving, storm-scale dynamics
- WoFSCast matches WoFS forecasts out to 2 hr and generates 6-hr ensemble forecasts in 30–40 s on 1 GPU
- Despite using an MSE loss, no substantial loss of information at smaller scales occurs out to 2 hr

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

M. L. Flora,
monte.flora@noaa.gov

Citation:

Flora, M. L., & Potvin, C. (2025). WoFSCast: A machine learning model for predicting thunderstorms at watch-to-warning scales. *Geophysical Research Letters*, *52*, e2024GL112383. https://doi.org/10.1029/2024GL112383

Received 4 SEP 2024
Accepted 4 MAR 2025

Abstract Developing AI models that match or exceed the forecast skill of numerical weather prediction (NWP) systems but run much more quickly is a burgeoning area of research. Most AI-NWP models, however, have been trained on global ECMWF Reanalysis version 5 data, which does not resolve storm-scale evolution. We have therefore adapted Google's GraphCast framework for limited-area, storm-scale domains, then trained on archived forecasts from the Warn-on-Forecast System (WoFS), a convection-allowing ensemble with 5-min forecast output. We evaluate the WoFSCast predictions using object-based verification, grid-based verification, spatial storm structure assessments, and spectra analysis. The WoFSCast closely emulates the WoFS environment fields, matches 70%–80% of WoFS storms out to 2-hr forecast times, and suffers only modest blurring. When verified against observed storms, WoFSCast produces contingency table statistics and fractions skill scores similar to WoFS. WoFSCast demonstrates that AI-NWP can be extended to rapidly evolving, small-scale phenomena like thunderstorms.

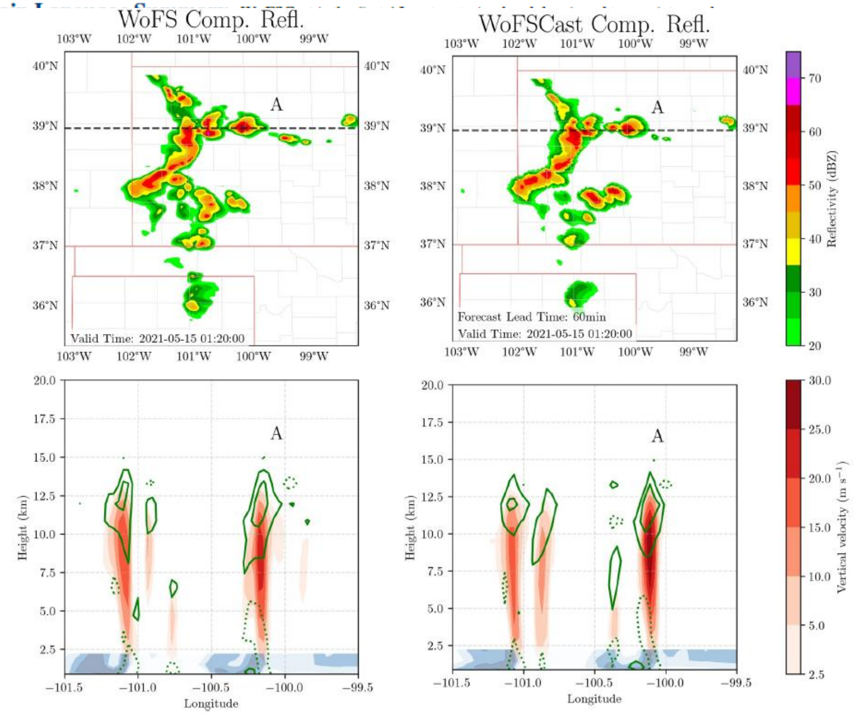


Figure 4. Top row: (left) WoFS member 9 composite reflectivity valid at 01:20 UTC on 15 May 2021 (initialized 2 hr earlier) and (right) 60-min WoFSCast composite reflectivity forecast valid at the same time. Bottom row: Vertical cross-section along the dashed line shown in the top row. Red color fill represents vertical velocity, with divergence in solid green contours and convergence in dashed green contours. The perturbed virtual potential temperature in the lowest 2.5 km indicates cold pool density (blue color fill). Values for divergence and virtual potential temperature were omitted for clarity. "A" highlights the supercell discussed in the text.

Kilometer-Scale Convection Allowing Model Emulation using Generative Diffusion Modeling

Jaideep Pathak*[1], Yair Cohen*[1], Piyush Garg*[1], Peter Harrington*[2], Noah Brenowitz[1], Dale Durran[1,3], Morteza Mardani[1], Arash Vahdat[1], Shaoming Xu†[4], Karthik Kashinath[1], Michael Pritchard[1]

NVIDEA StormCast –2024

August 18, 2024

[1]NVIDIA Corporation
[2]Lawrence Berkeley National Laboratory
[3]University of Washington
[4]University of Minnesota
*Equal Contribution

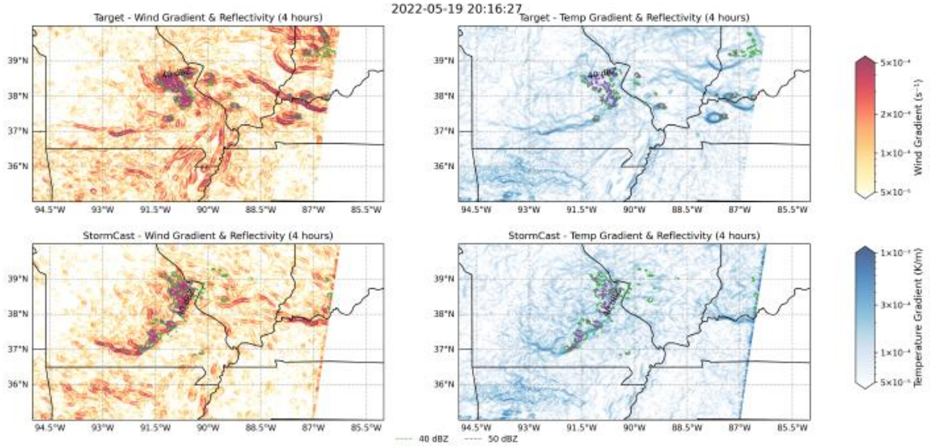
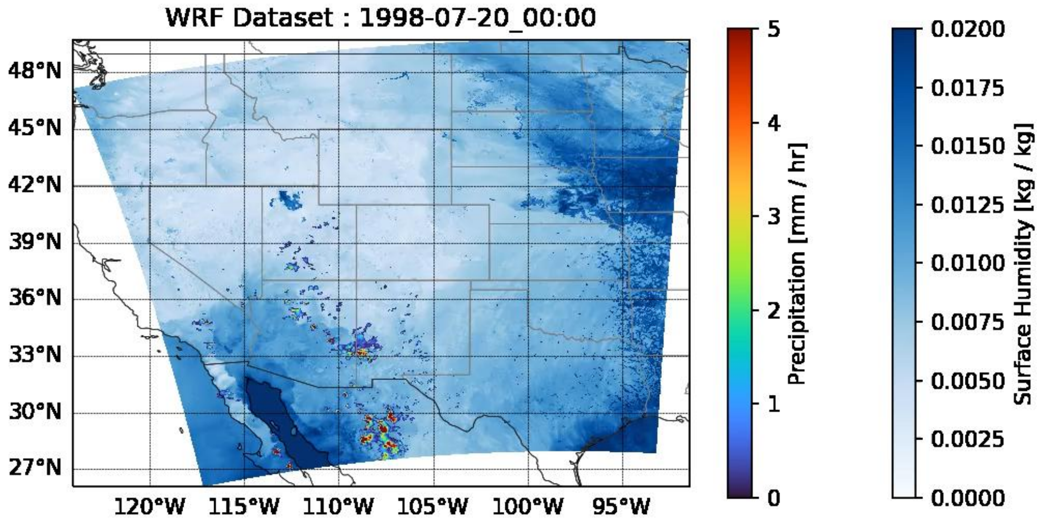


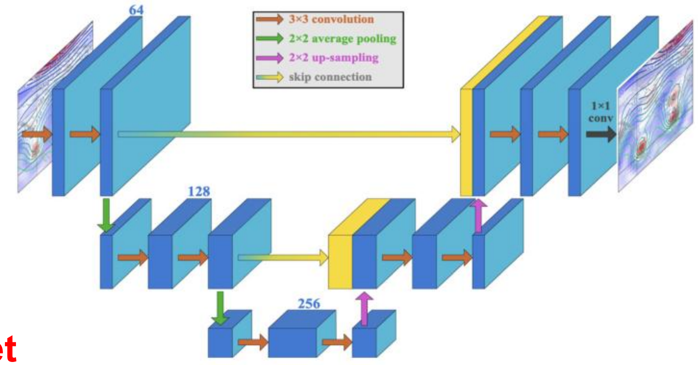
Figure 7: Representative fine-scale features of the horizontal gradients of 10m horizontal wind (left column) and 125m air temperature gradient (right column) for target (top rows) and StormCast 4-hour forecasts (bottom rows), suggestive of cold pool related gust fronts. This case was initialized on 2022-05-19 17:30:00 UTC. Composite radar reflectivity contours at 40 and 50 dBZ are overlain in green and purple.

NSF NCAR Digital Twins

Mesoscale

Gutmann, Sha, McGinnis, McCrary, Newman, Smith,…

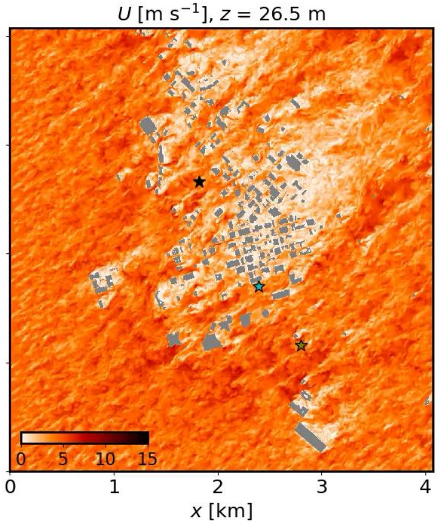




UNet

Urban Scale

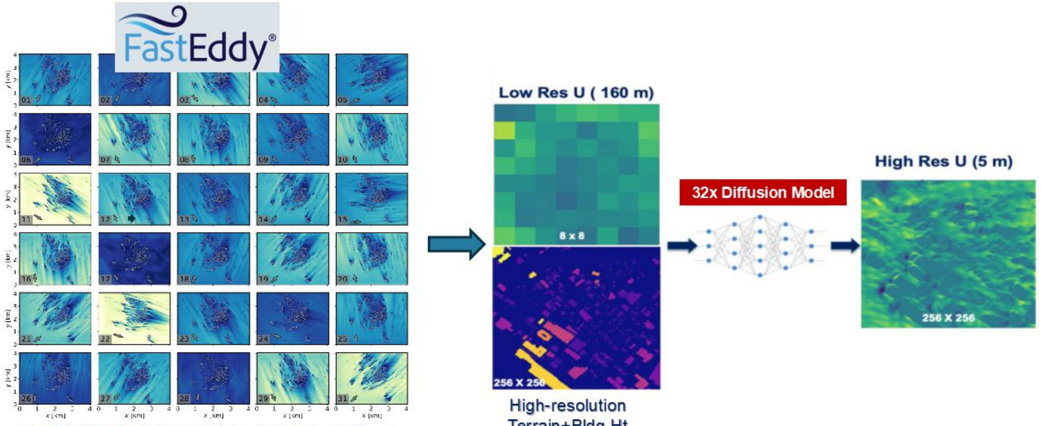
Haupt, Dettling, Brummet, Sha, Kosovic, Boenert,…



LES: 5-m resolution

Microscale Downscaling – Urban Scale - Methodology

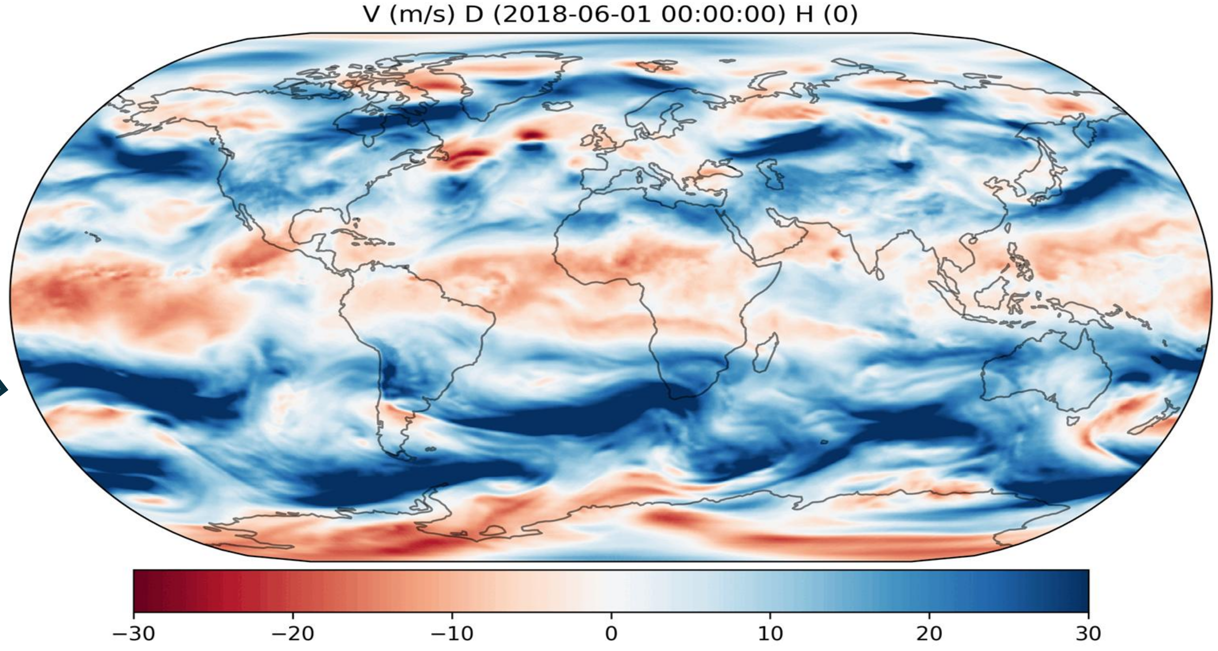
Diffusion Model + Neural Operator



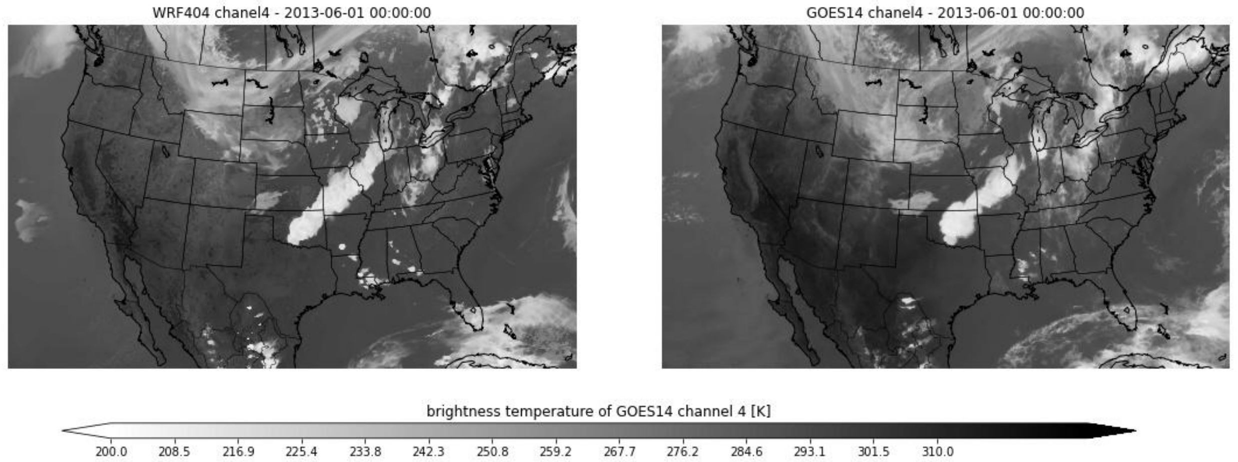
FastEddy Training Data - 5 m resolution
Dallas/Fort Worth – 30 days – noon hour
Using 13 days

Bringing it Together for New Generative Platform

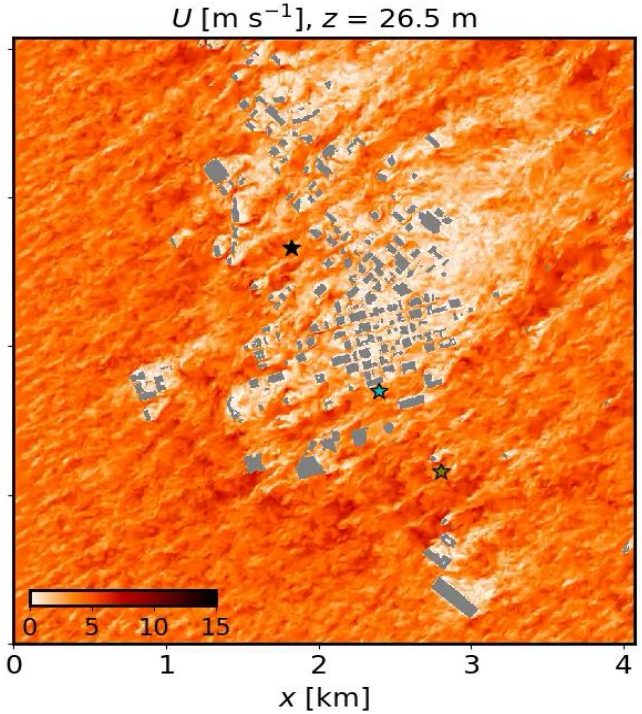
CISL Team: John Schreck (CISL), Will Chapman (CGD), DJ Gagne (CISL)



Mesoscale Applications
CONUS 404 Training Data



Microscale Applications
FastEddy Training Data



Summary:

- Machine Learning has become a necessary component of modern applications in weather forecasting.
- AIWP is revolutionizing weather prediction.
- It’s all very new and we’ll need to understand the full picture.
- The prospect for the future is bright.