Community Research Earth Digital Intelligence Twin (CREDIT)

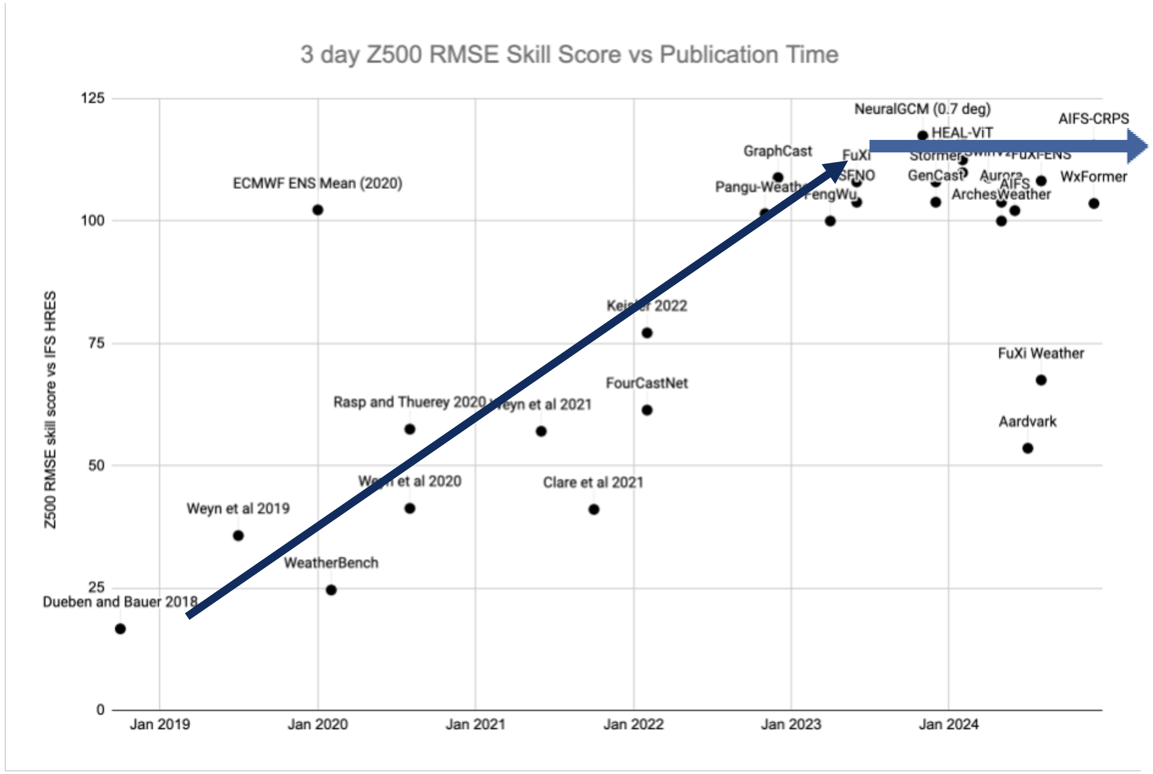
David John Gagne
Machine Learning Scientist II, US NSF NCAR

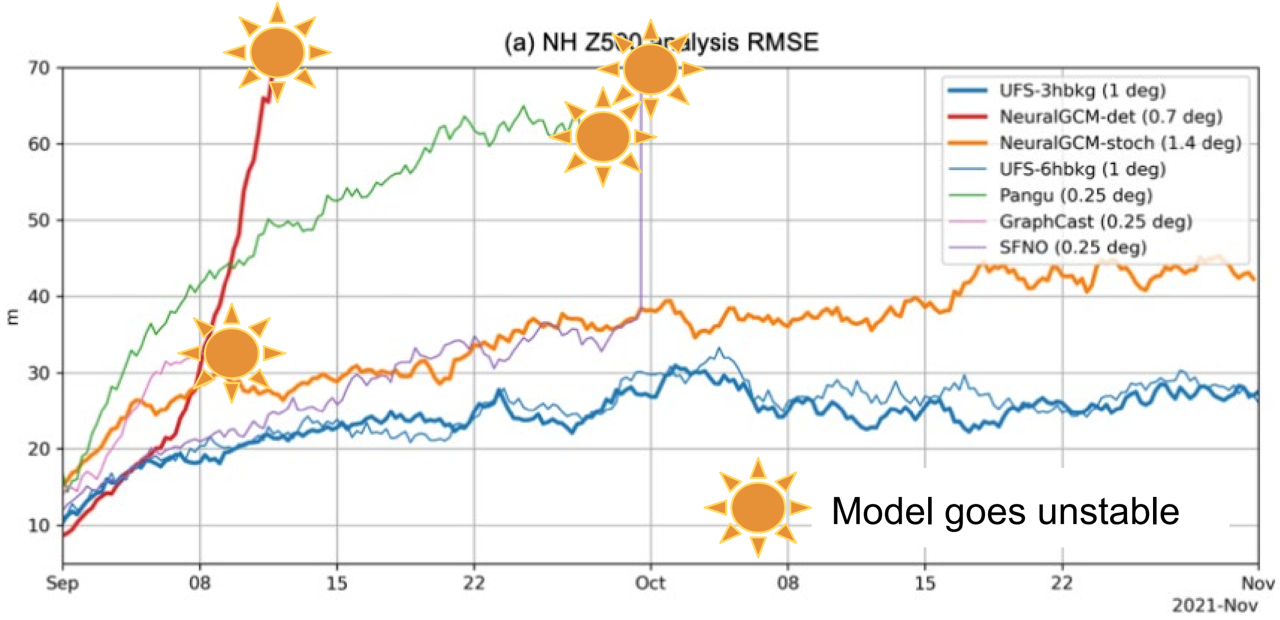
CREDIT Team: *John Schreck, Will Chapman, Kyle Sha, David John Gagne, Dhamma Kimpara, Arnold Kazadi, Seth McGinnis, Negin Sobhani, Ben Kirk, Judith Berner, Charlie Becker, Gabrielle Gantos, Kirsten Mayer, Laure Zanna*

June 4, 2025

This material is based upon work supported by the NSF National Center for Atmospheric Research, which is a major facility sponsored by the U.S. National Science Foundation under Cooperative Agreement No. 1852977.

Motivation





Source: Slivinski et al. 2024

After 4 years of rapid advancement in accuracy, further advancements in AI weather modeling have shown diminishing returns in improving global metrics.

Experiments with data assimilation and ensembles have revealed physical inconsistencies and instabilities that require more engagement with the data and physics.

Our 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

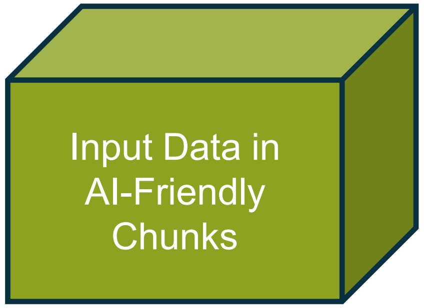
Models

- WXFormer
- NCAR FuXi
- Graph Recurrent Transformer

Physics

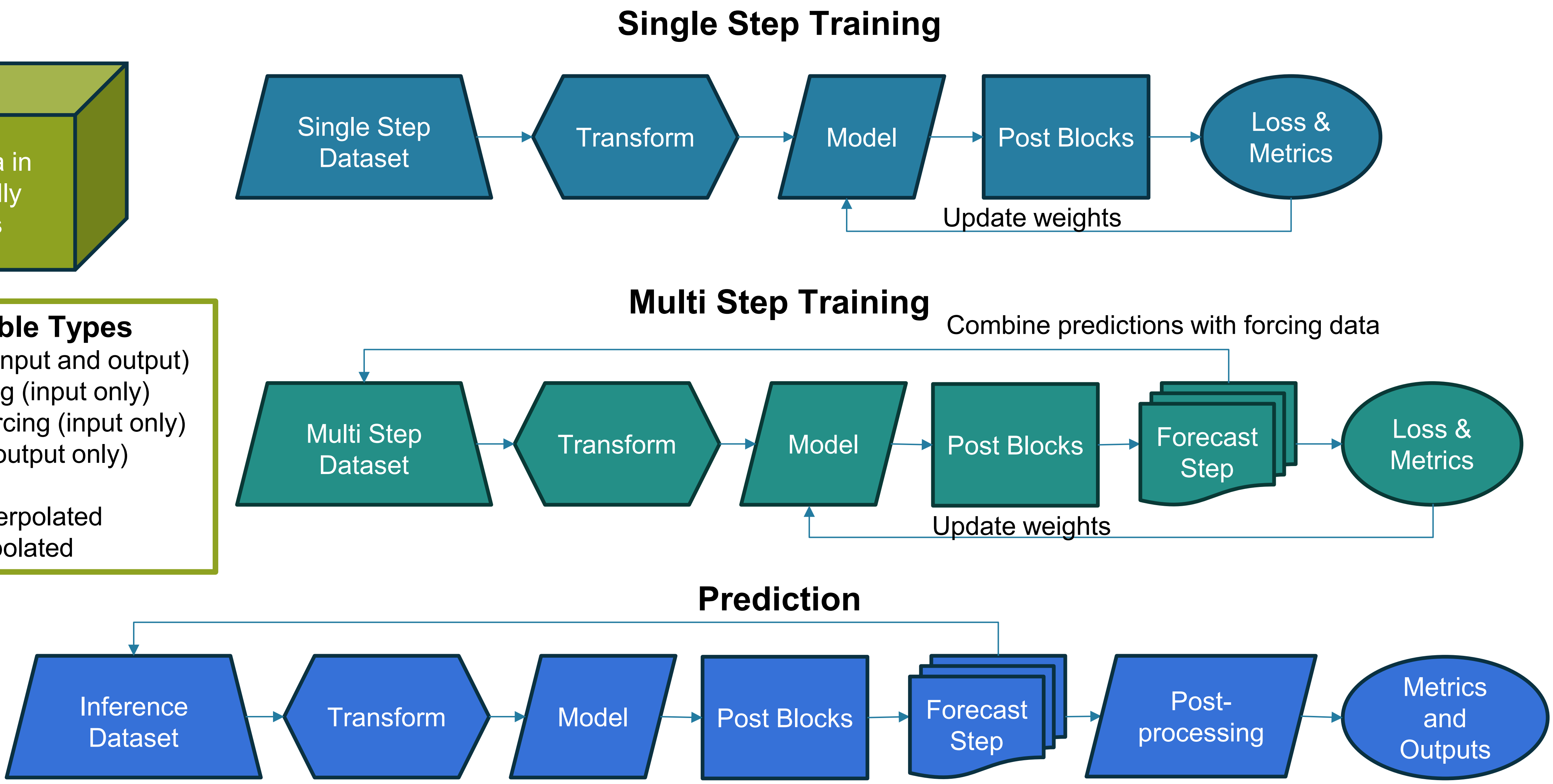
- Mass Conservation
- Energy Conservation
- Moisture Budget

CREDIT Components



Variable Types

Prognostic (input and output)
Static Forcing (input only)
Dynamic Forcing (input only)
Diagnostic (output only)
Derived
Pressure interpolated
Height interpolated



CREDIT WXFormer v1: Training Data

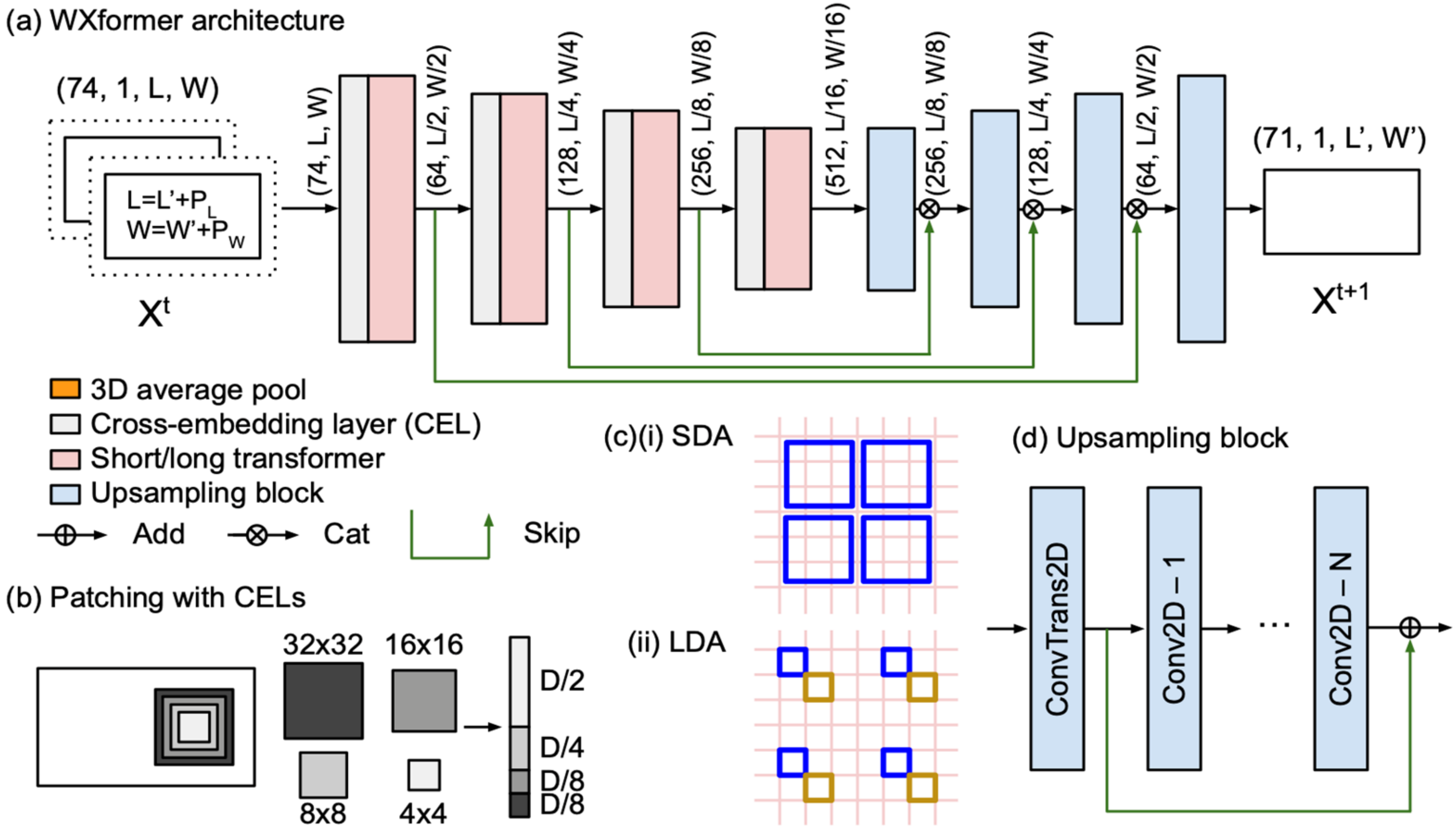
- ERA5 hybrid sigma-pressure level data on 0.28 deg. Gaussian grid (1280 x 640 grid cells)
  - 1979-2014 training
  - 2014-2017 validation
  - 2018-2022 testing
- State variables on 16 hybrid-sigma levels sampled from the 137 ERA5 levels
- Surface and 500 hPa variables
- Forcing
  - Integrated solar irradiance at TOA
  - Land Sea Mask
  - Geopotential at surface

Table 1: Input Variables and Their Units

Type	Variable Name	Short Name	Units	Usage
Model level variable	Zonal Wind	U	$\mathrm{m \cdot s^{-1}}$	Prognostic
Model level variable	Meridional Wind	V	$\mathrm{m \cdot s^{-1}}$	Prognostic
Model level variable	Air Temperature	T	K	Prognostic
Model level variable	Specific Humidity	Q	$\mathrm{kg \cdot kg^{-1}}$	Prognostic
Single level variable	Surface Pressure	SP	Pa	Prognostic
Single level variable	2-Meter Temperature	t2m	K	Prognostic
Single level variable	Meridional Wind at 500 hPa	V500	$\mathrm{m \cdot s^{-1}}$	Prognostic
Single level variable	Zonal Wind at 500 hPa	U500	$\mathrm{m \cdot s^{-1}}$	Prognostic
Single level variable	Temperature at 500 hPa	T500	K	Prognostic
Single level variable	Geopotential Height at 500 hPa	Z500	m	Prognostic
Single level variable	Specific Humidity at 500 hPa	Q500	$\mathrm{kg \cdot kg^{-1}}$	Prognostic
Invariant variable	Geopotential at surface	Z_{SFC}	$\mathrm{m^2 \cdot s^{-2}}$	Input-only
Invariant variable	Land Sea Mask	LSM	n/a	Input-only
Forcing variable	Integrated instantaneous solar irradiance	I_s	$\mathrm{J \cdot m^{-2}}$	Input-only

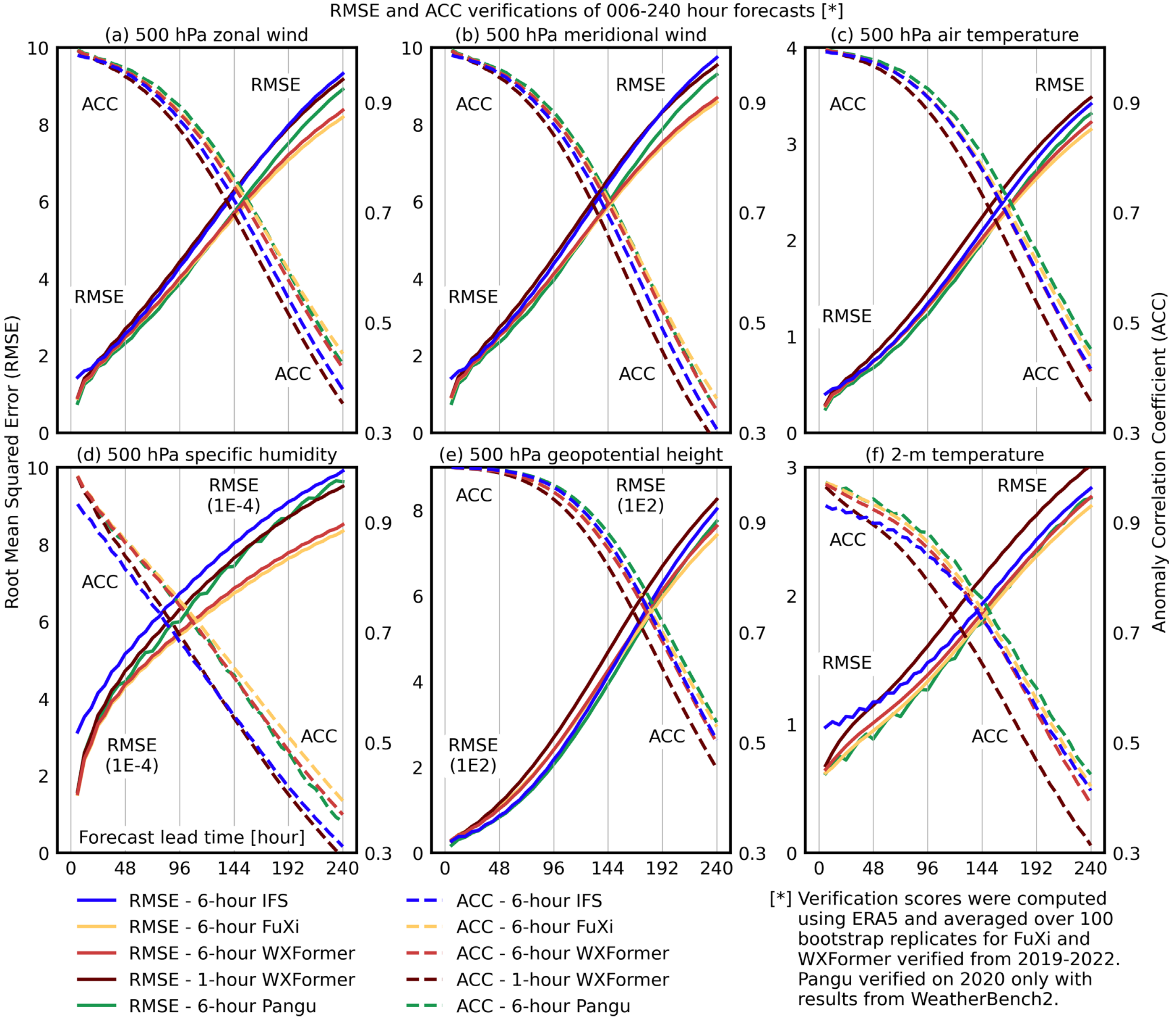
CREDIT WXFormer Model Architecture

Spectral normalization of weights during training
- prevent exploding and vanishing gradients
- improve model stability



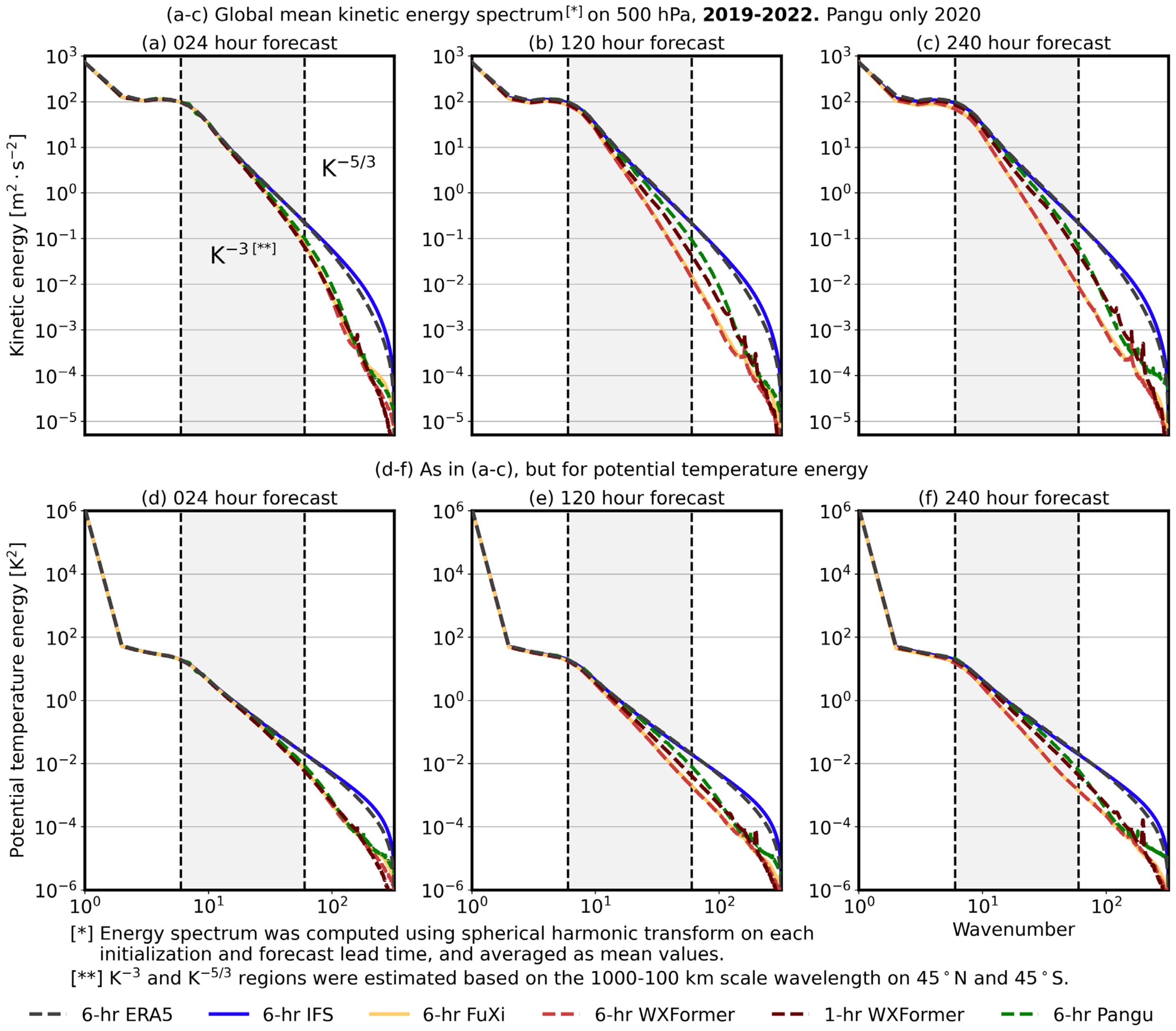
Global Verification

- Both WXFormer and MILES FuXi are outperforming IFS for all surface variables
- Larger gains with specific humidity and surface temperature
- Bigger gains at longer lead times
- Performance consistent with other AI NWP models
- Competitive with Pangu-Weather but more stable



[*] Verification scores were computed using ERA5 and averaged over 100 bootstrap replicates for FuXi and WXFormer verified from 2019-2022. Pangu verified on 2020 only with results from WeatherBench2.

Kinetic Energy Spectra



Variations by Vertical Level

WXFormer 1 Hour

ERA5

Physics Conservation (Work led by Kyle Sha)

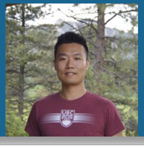
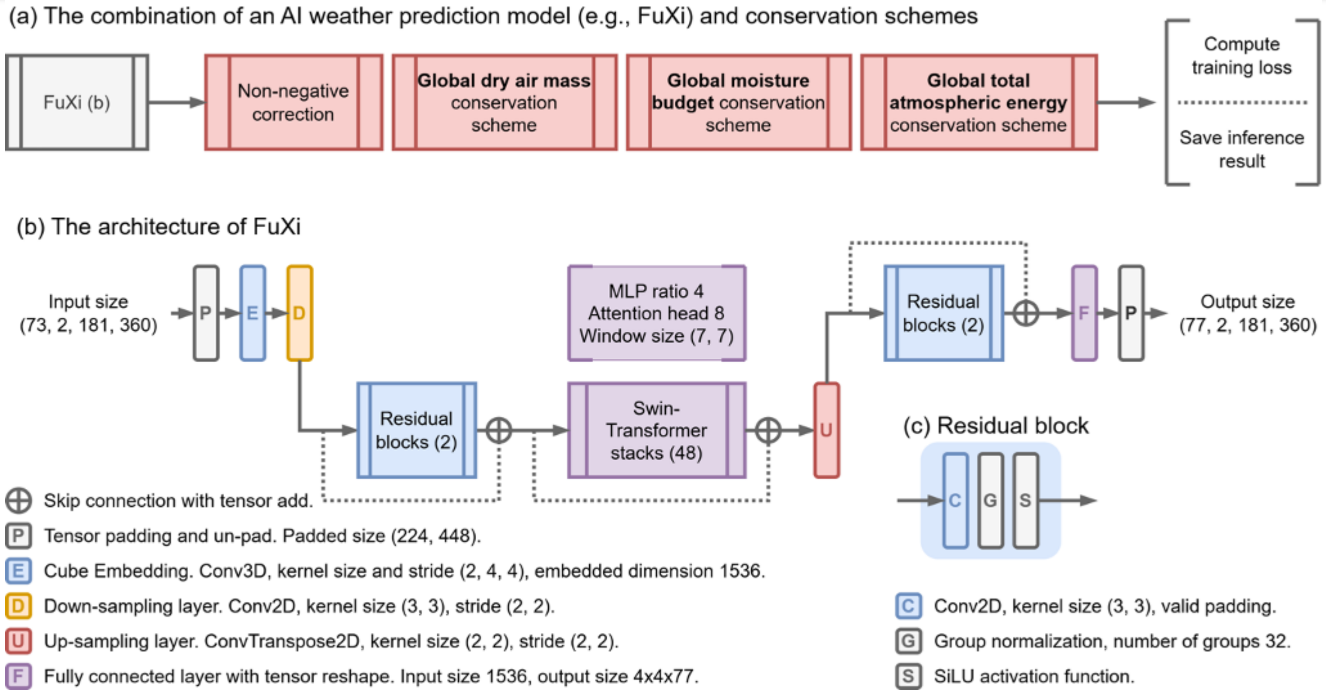


Table 1. The variables of interest in this study.

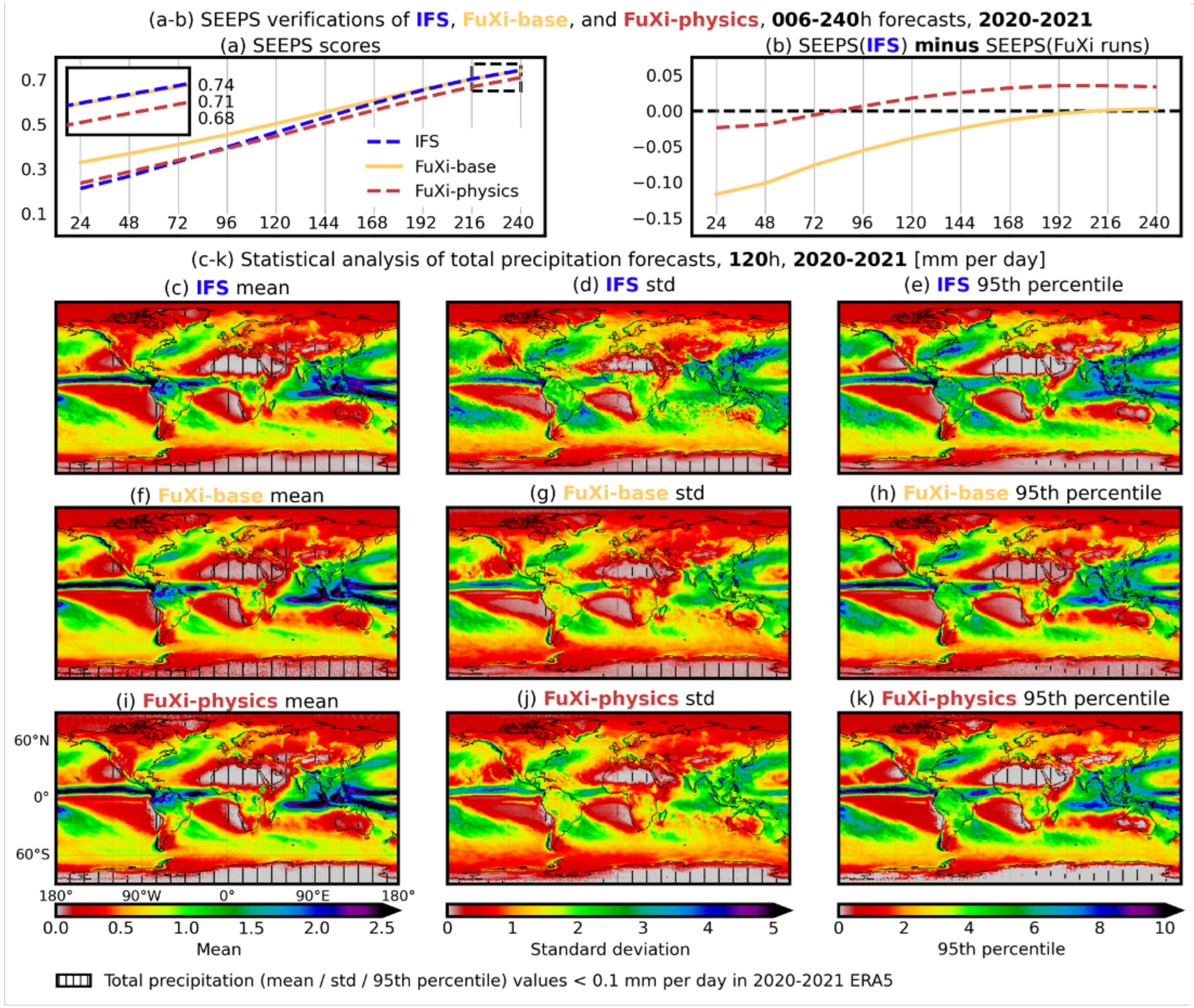
Type	Variable Name	Units	Role
Pressure level	Zonal Wind	$m \cdot s^{-1}$	Prognostic, Instantaneous
	Meridional Wind	$m \cdot s^{-1}$	
	Air Temperature	K	
	Specific Total Water[a]	$kg \cdot kg^{-1}$	
	Geopotential height	m	
Single level	Mean Sea Level Pressure	Pa	Prognostic, Instantaneous
	2-Meter Temperature	K	
	10-Meter Zonal Wind	$m \cdot s^{-1}$	
	10-Meter Meridional Wind	$m \cdot s^{-1}$	
Flux form[b]	Total Precipitation	m	Diagnostic, Cumulative
	Evaporation	m	
	Top-of-atmosphere Net Solar Radiation	$J \cdot m^{-2}$	
	Outgoing Longwave Radiation	$J \cdot m^{-2}$	
	Surface Net Solar Radiation	$J \cdot m^{-2}$	
	Surface Net Longwave Radiation	$J \cdot m^{-2}$	
	Surface Net Sensible Heat Flux	$J \cdot m^{-2}$	
	Surface Net Latent Heat Flux	$J \cdot m^{-2}$	
	Top-of-atmosphere Incident Solar Radiation	$J \cdot m^{-2}$	Input-only, Cumulative
Others	Sea-ice Cover	n/a	Input-only, Instantaneous
	Geopotential at the Surface	$m^2 \cdot s^{-2}$	Input-only, Static
	Land-sea Mask	n/a	Input-only, Static
	Soil Type	n/a	Input-only, Static

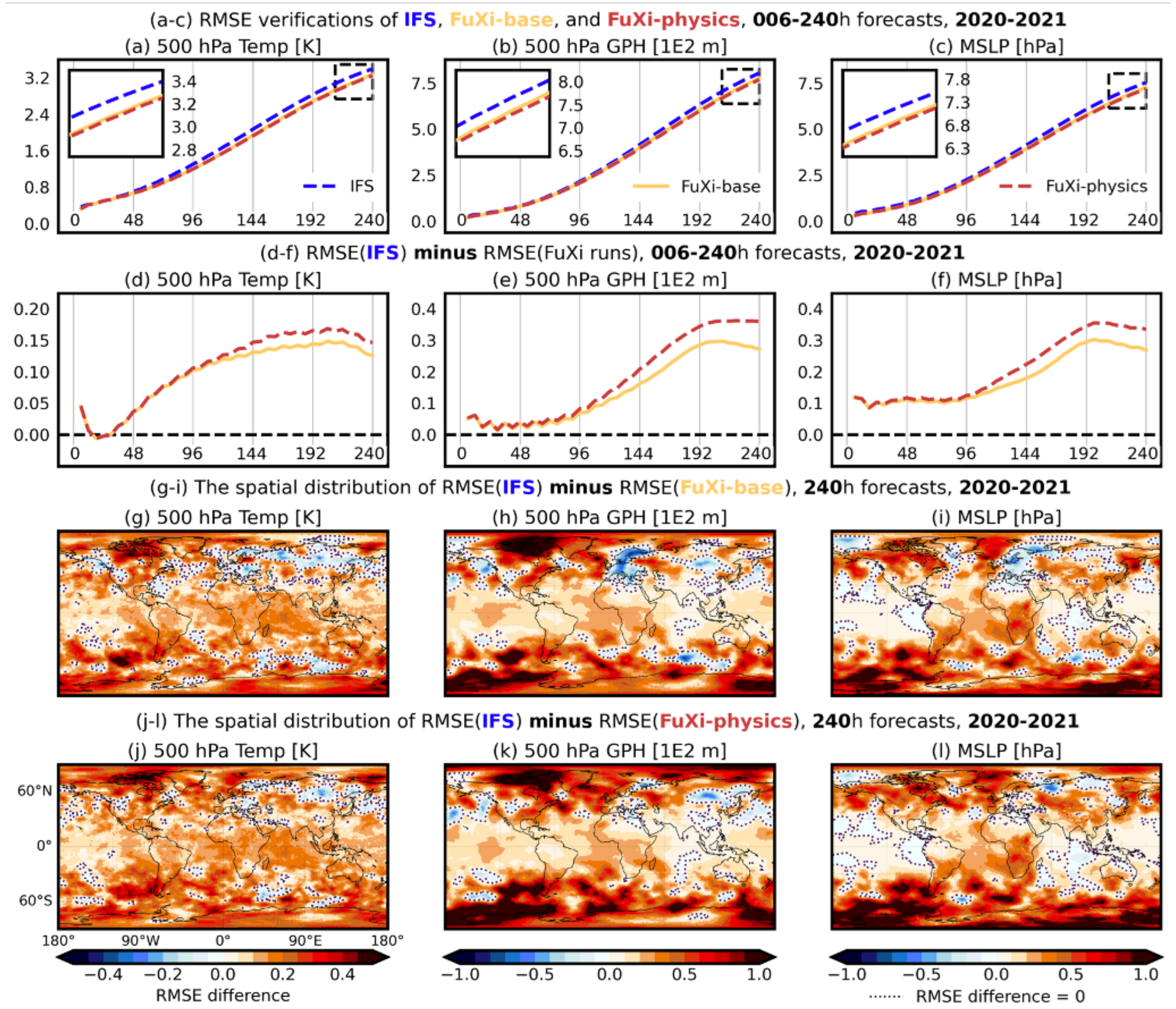
[a] Specific total water is the combination of specific humidity, cloud liquid water content, and rainwater content.
[b] Flux form variables are accumulated every 6 hours. Downward flux is positive.



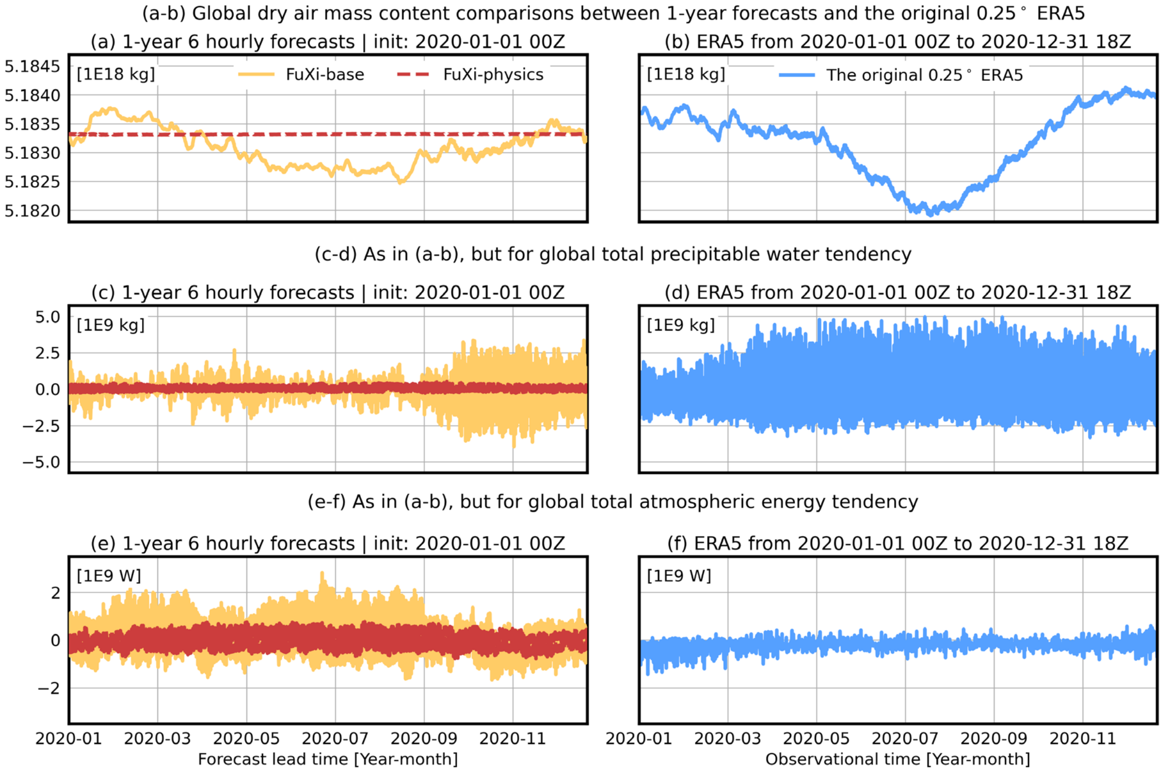
Data: ERA5 conservatively regridded to 1 degree
Loss: Latitude-weighted MSE
Ingredients for Physics Constraints in AI Weather Prediction
1. Sufficient variables to calculate mass, moisture, and energy budgets
2. Conservation layers that adjust data to conserve mass, moisture, and total energy across the globe to match initial values with multiplicative scaling

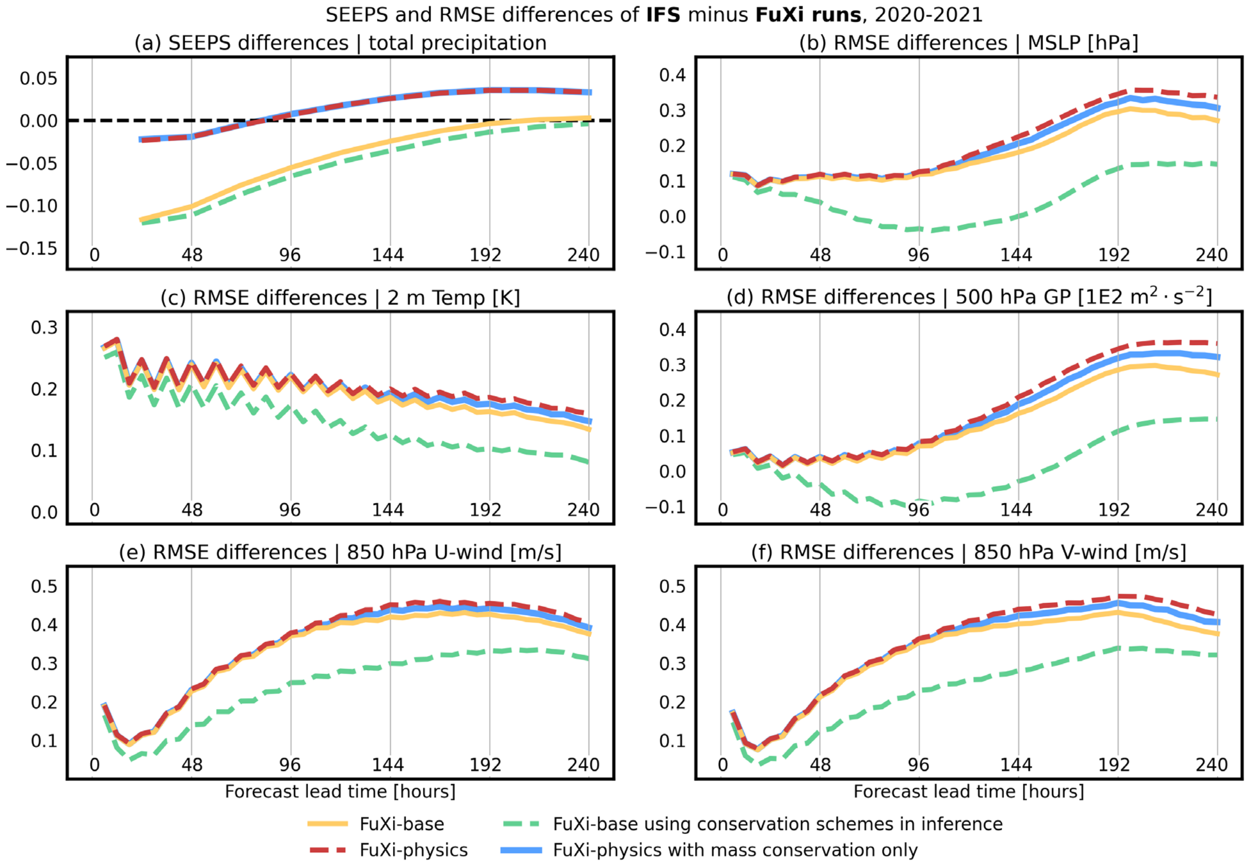
Forecast Improvements





Ablation Analysis

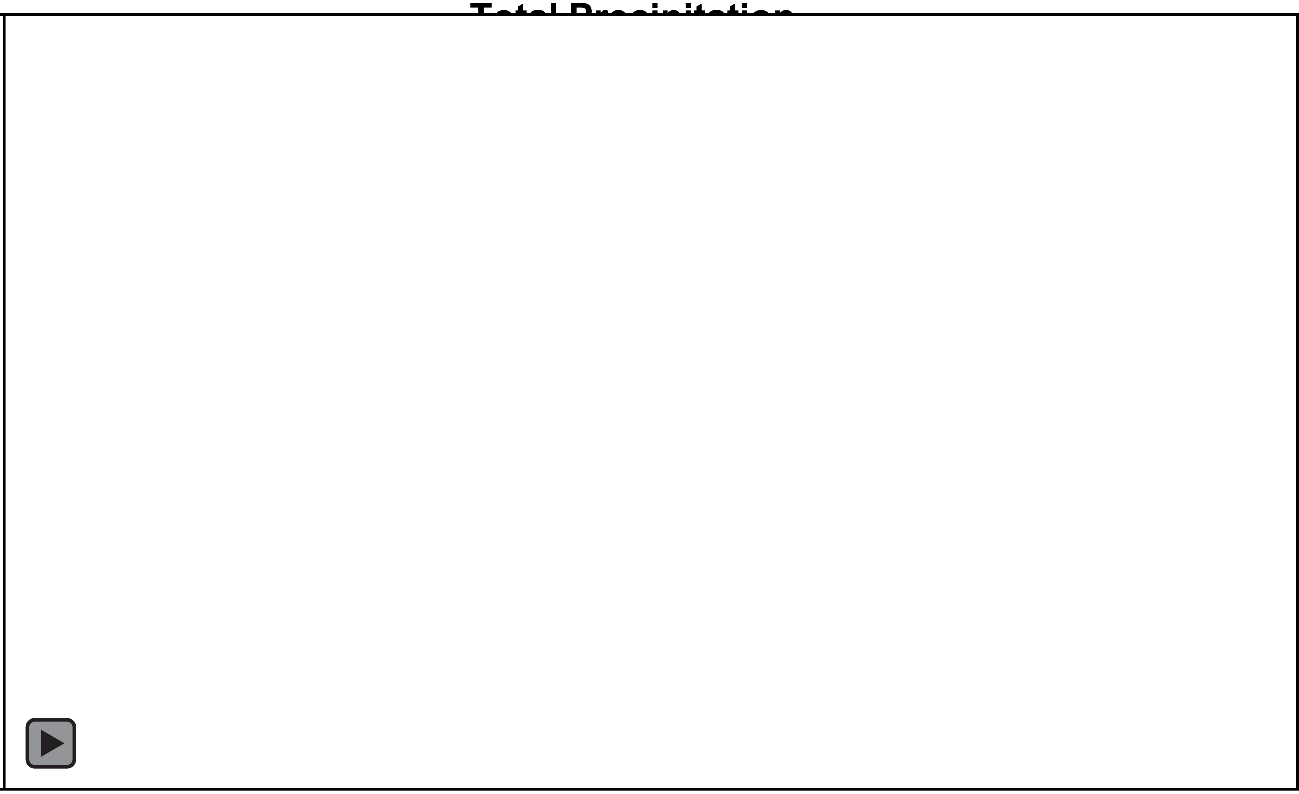




Physics-Constrained Case Example

Top of atmosphere net thermal radiation

Total Precipitation



CREDIT User Stories

Graduate Student

Goal: Wants to investigate the effect of changes in sea surface temperature patterns on snow depth in the Andes.

With CREDIT: Starts with a pre-trained CREDIT model, generates ensemble forecasts for 5 years on one Derecho node in a day and applies ML winter precipitation type algorithm to partition rain and snow and estimate snow depth. Adjusts SST distribution and re-runs ensemble the next day.

Without CREDIT: Grad student spends a month generating MPAS ensemble runs and then another month writing code to calculate p-type and parse through the data.

Risk Modeler

Goal: Simulate the compound risk of heat waves with deadly wet bulb globe temperatures following tropical cyclones that reduce electricity and A/C.

With CREDIT: Run 50 initializations of CAMulator for 100 years each on 1 GPU for 1 day on an AWS GPU node. Run tropical cyclone storm tracker and heat anomaly algorithm to find collocated tropical cyclone/heat wave events in parallel with each ensemble member and only save data around combined events. Perform analytics to assess global risks the next day.

Without CREDIT: Requests for 1 million core-hours on Derecho. Run for 357 compute-days across 500 nodes to generate regular CAM ensemble in a week. Then spend month writing parallel analytics script to parse through many terabytes of model output.

CREDIT Future Directions

Open Questions

- Ensemble generation: what is the most accurate method with least latency?
- Tradeoffs between data volume, model size, input data size, and types of physical constraints
- How to improve vertical exchange of information in model, especially between troposphere and stratosphere
- End-to-end black box model vs more interpretable/tunable collection of component models?

Next Steps

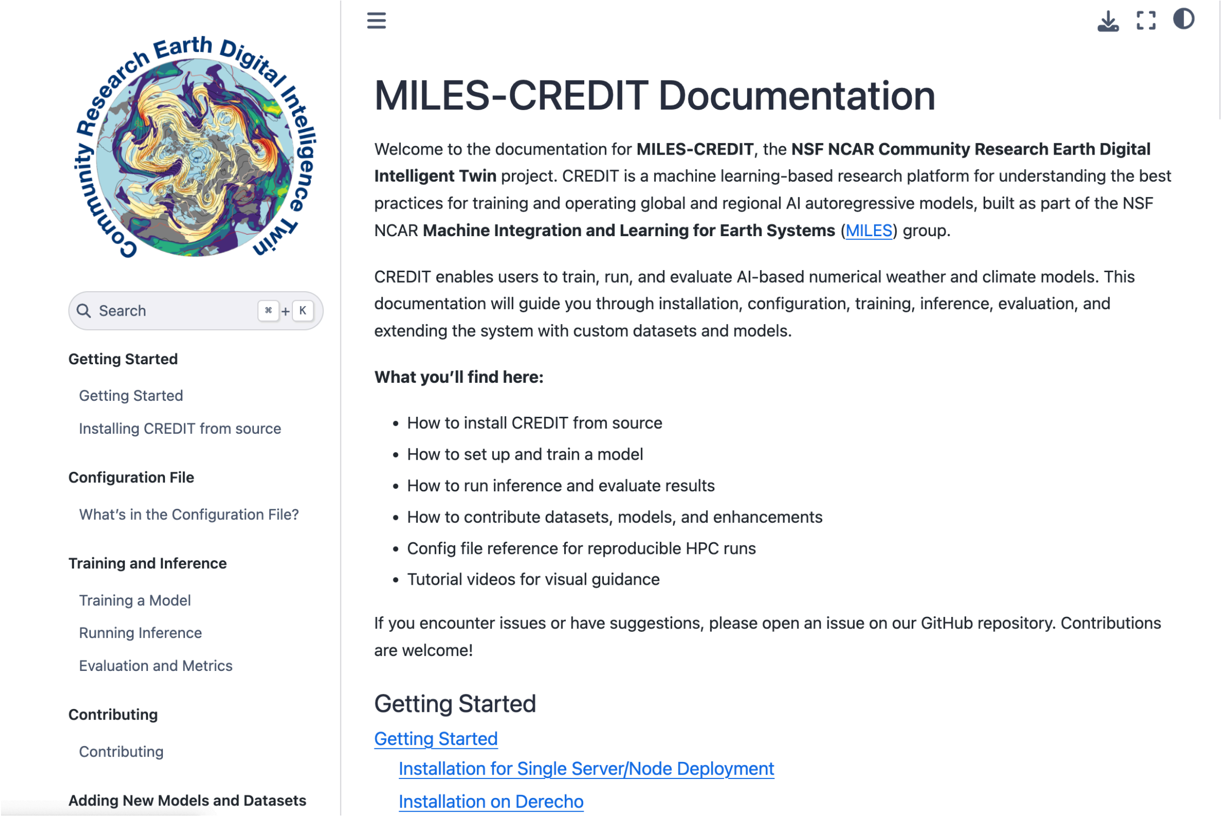
- Improve usability of CREDIT with software engineering support
- Adding ensemble generation
- Regional model training and evaluation
- S2S and longer scale rollout evaluation
- Training a new weather model with more vertical levels at 0.25 degree or finer resolution

Summary

- CREDIT opens a new pathway to customization of the whole AI weather and climate modeling pipeline
- Physics constraints and data choices greatly improve model realism
- Paper accepted in npj Climate and Atmospheric Science!
- CREDIT source code and models: https://github.com/NCAR/miles-credit
- Links to CREDIT papers: https://miles.ucar.edu/projects/credit/

Contact Me
Email: dgagne@ucar.edu
Bluesky: @DJGagneDos
Github: djgagne

Version 2025.2.0 is out now!


Community Research Earth Digital Intelligence Twin

Search

Getting Started
Getting Started
Installing CREDIT from source

Configuration File
What's in the Configuration File?

Training and Inference
Training a Model
Running Inference
Evaluation and Metrics

Contributing
Contributing

Adding New Models and Datasets

MILES-CREDIT Documentation

Welcome to the documentation for **MILES-CREDIT**, the **NSF NCAR Community Research Earth Digital Intelligent Twin** project. CREDIT is a machine learning-based research platform for understanding the best practices for training and operating global and regional AI autoregressive models, built as part of the NSF NCAR **Machine Integration and Learning for Earth Systems** (MILES) group.

CREDIT enables users to train, run, and evaluate AI-based numerical weather and climate models. This documentation will guide you through installation, configuration, training, inference, evaluation, and extending the system with custom datasets and models.

What you'll find here:

- How to install CREDIT from source
- How to set up and train a model
- How to run inference and evaluate results
- How to contribute datasets, models, and enhancements
- Config file reference for reproducible HPC runs
- Tutorial videos for visual guidance

If you encounter issues or have suggestions, please open an issue on our GitHub repository. Contributions are welcome!

Getting Started

Getting Started
- Installation for Single Server/Node Deployment
- Installation on Derecho