CAMulator: Fast Emulation of the Community Atmosphere Model

Will Chapman

Project Scientist , US NSF NCAR

CREDIT Team: *John Schreck, David John Gagne, 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 11, 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.

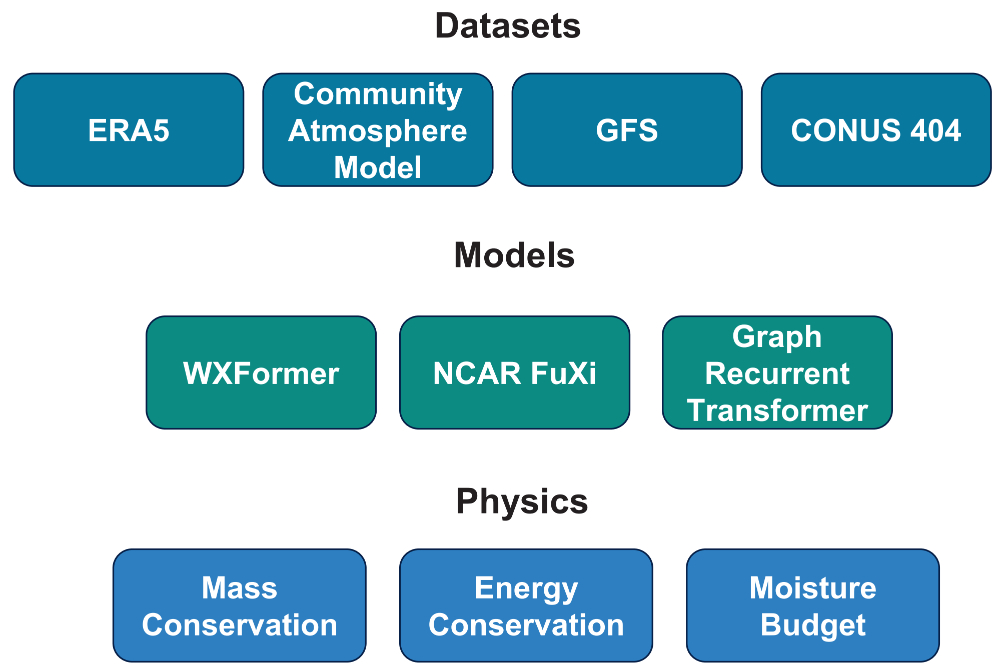
Our Framework: CREDIT

What is CREDIT?

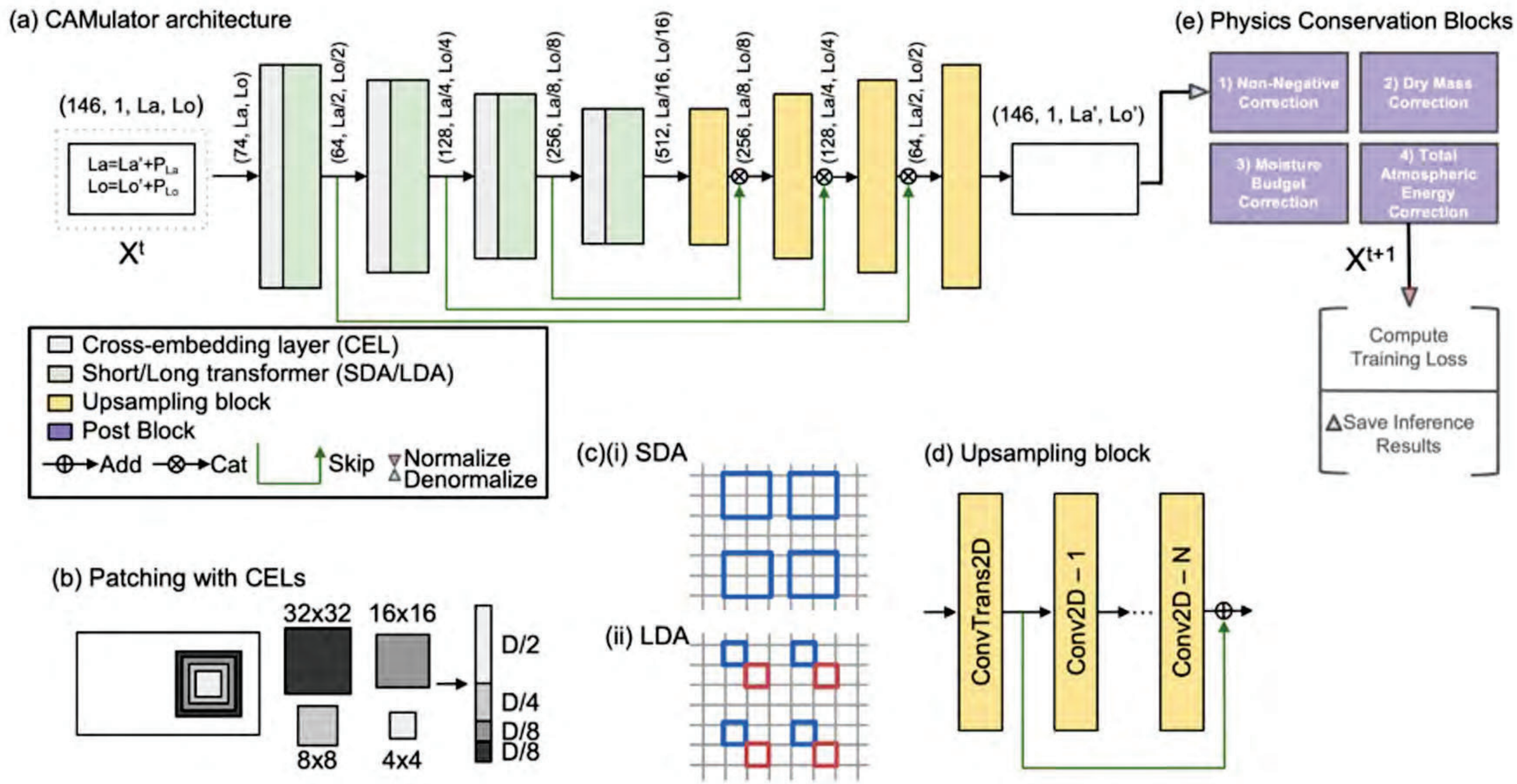
An open foundational platform for developing and deploying AI weather and Earth system prediction models.

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.



Variables and Architecture



- Trained on CAM over 30 year observed record forcing
- Throughput of 480 Sim. Years/Day

Chapman, Will, et al. "CAMulator: Fast Emulation of the Community Atmosphere … "*arXiv preprint arXiv:*2504.06007 (2025).

FHIST Run

WE ARE GOING TO EMULATE THESE VARIABLES:

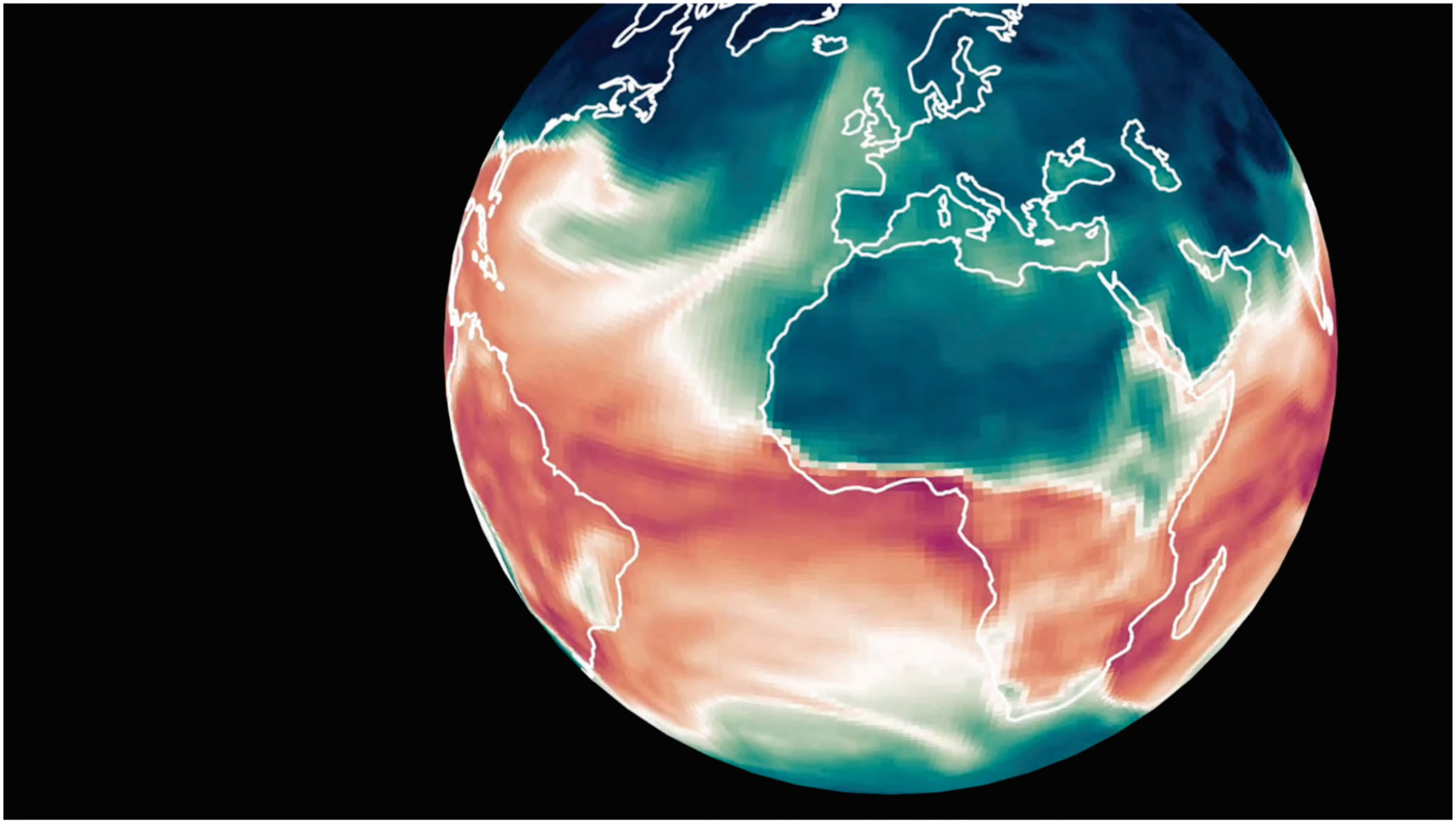
Variable	Description	Units	Single Level/Levels	I/O
Prognostic Variables (Input and Output)				
U	Zonal Wind	m/s	32 levels	Input/Output
V	Meridional Wind	m/s	32 levels	Input/Output
T	Temperature	K	32 levels	Input/Output
Qtot	Specific Total Water	kg/kg	32 levels	Input/Output
Diagnostic Variables (Output Only)				
PRECT	Precipitation Rate	m	Single Level	Output
CLDTOT	Total Cloud Cover	fraction	Single Level	Output
CLDHGH	High Cloud Cover	fraction	Single Level	Output
CLDLOW	Low Cloud Cover	fraction	Single Level	Output
CLDMED	Medium Cloud Cover	fraction	Single Level	Output
TAUX	Zonal Wind Stress	N/m^2	Single Level	Output
TAUY	Meridional Wind Stress	N/m^2	Single Level	Output
U10	10m Wind Speed	m/s	Single Level	Output
QFLX	Surface Moisture Flux	m	Single Level	Output
FSNS	Net Solar Flux at Surface	J/m^2	Single Level	Output
FLNS	Net Longwave Flux at Surface	J/m^2	Single Level	Output
FSNT	Net Solar Flux at TOA	J/m^2	Single Level	Output
FLNT	Net Longwave Flux at TOA	J/m^2	Single Level	Output
SHFLX	Sensible Heat Flux	J/m^2	Single Level	Output
LHFLX	Latent Heat Flux	J/m^2	Single Level	Output
Surface Variables Prognostic (Input and Output)				
PS	Surface Pressure	Pa	Single Level	Input/Output
TREFHT	Near-Surface Air Temperature	K	Single Level	Input/Output
Dynamic Forcing Variables (Input Only)				
SOLIN	Incoming Solar Radiation	J/m^2	Single Level	Input
SST	Sea Surface Temperature	K	Single Level	Input
Static Forcing Variables (Input Only)				
Surface Geop.	Normalized Surface Height	m^2/s^2	Single Level	Input
Land-Sea Mask	Land Mask × Cosine Latitude	unitless	Single Level	Input

WE FORCE THE MODEL WITH THESE VARIABLES (AMIP)

Data Prep Workflow

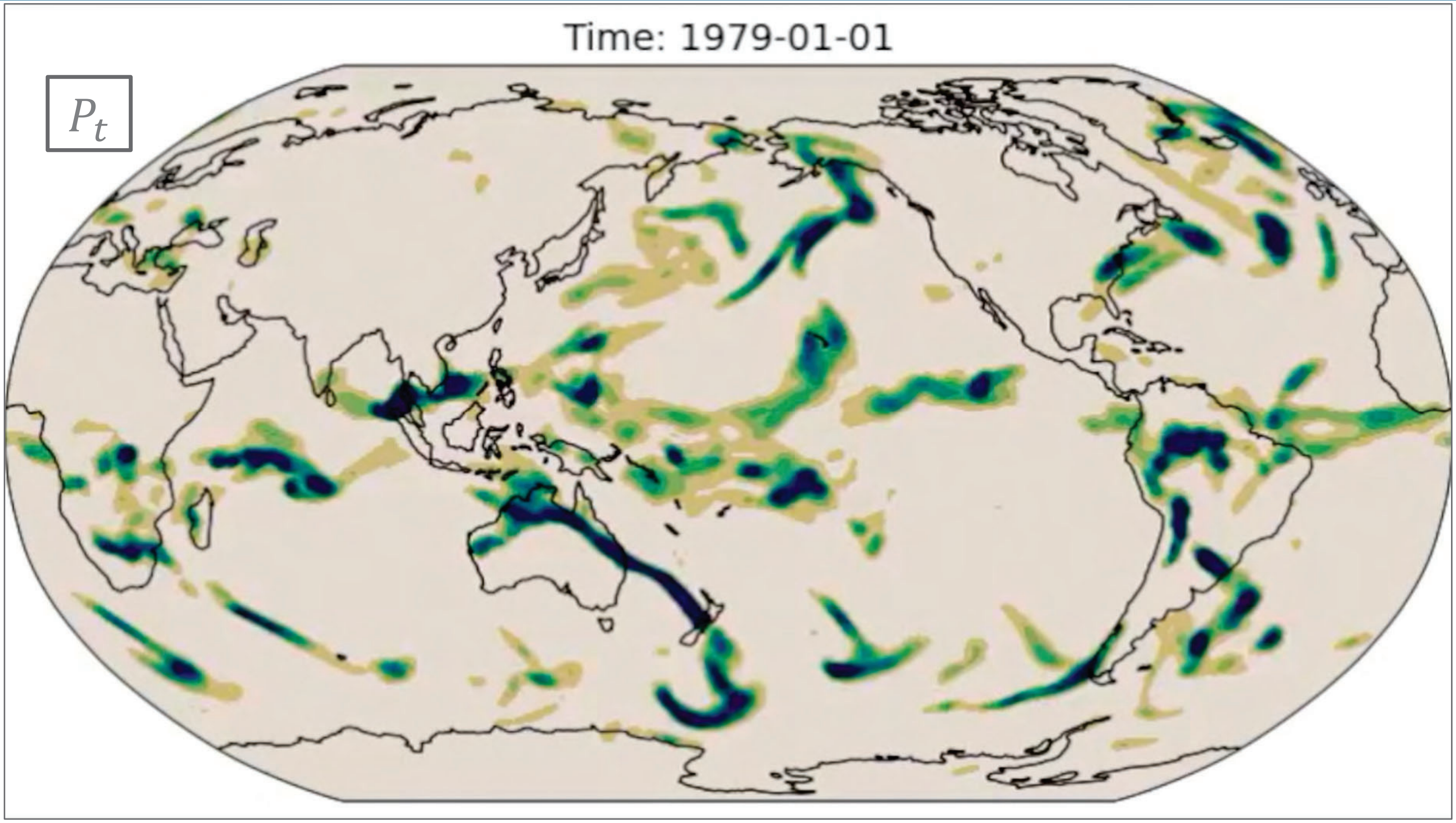
1) Conduct a 35-year FHIST run (1979-2014) of ***CAM6 in CESMv2.1.5***
2) Collect 6-hourly data
3) Compute Total Water + Convert to Flux forms
4) Gather Static & Forcing Data

CAMulator V1.0



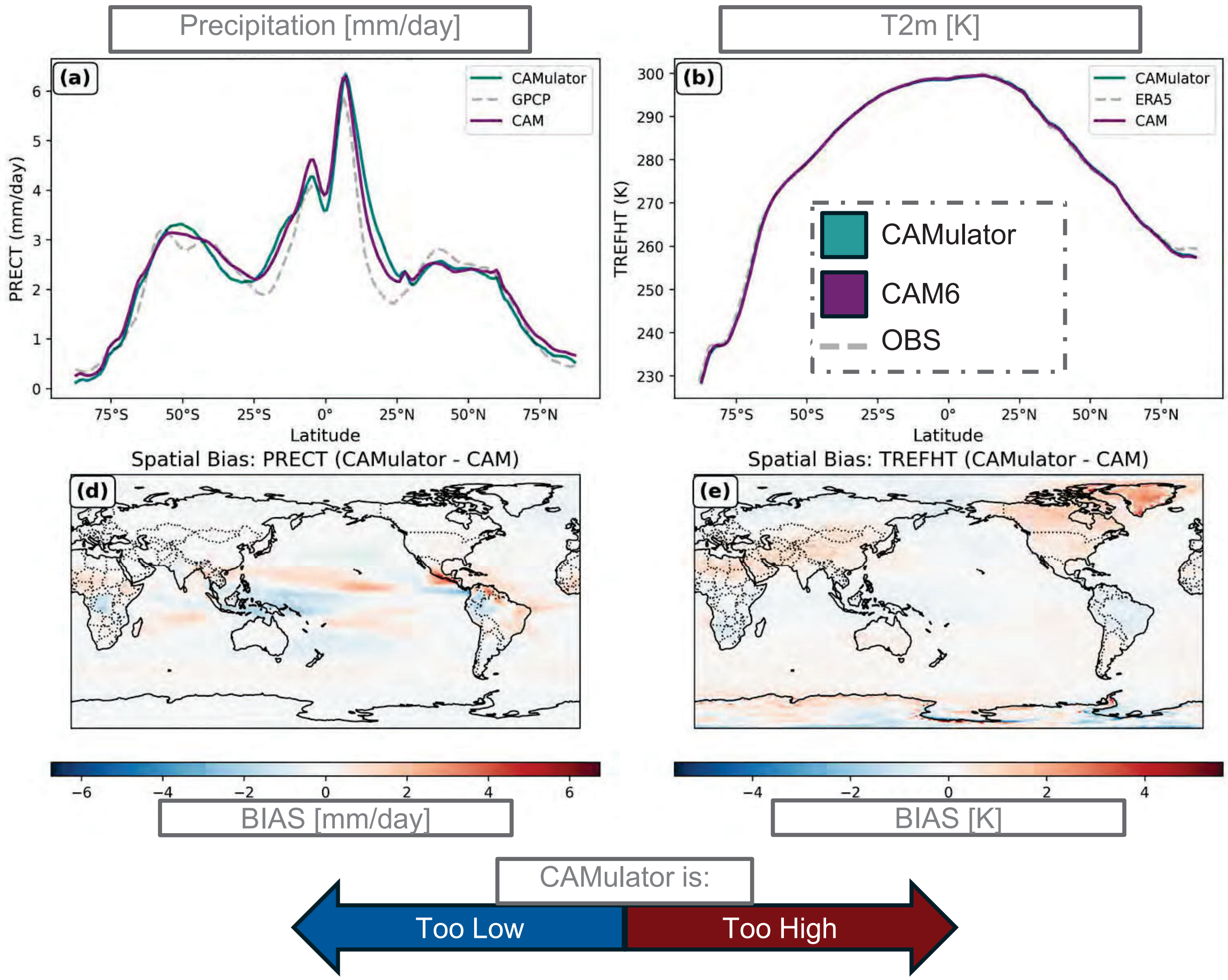
Chapman, Will, et al. "CAMulator: Fast Emulation of the Community Atmosphere … "*arXiv preprint arXiv:*2504.06007 (2025).

Precipitation Climatology V.1.0



$$P_t = \frac{1}{t}\sum_{i=1}^{t} PRECIPITATION(i)$$

Climatological Mean Bias Error: CAMulator V.1.0



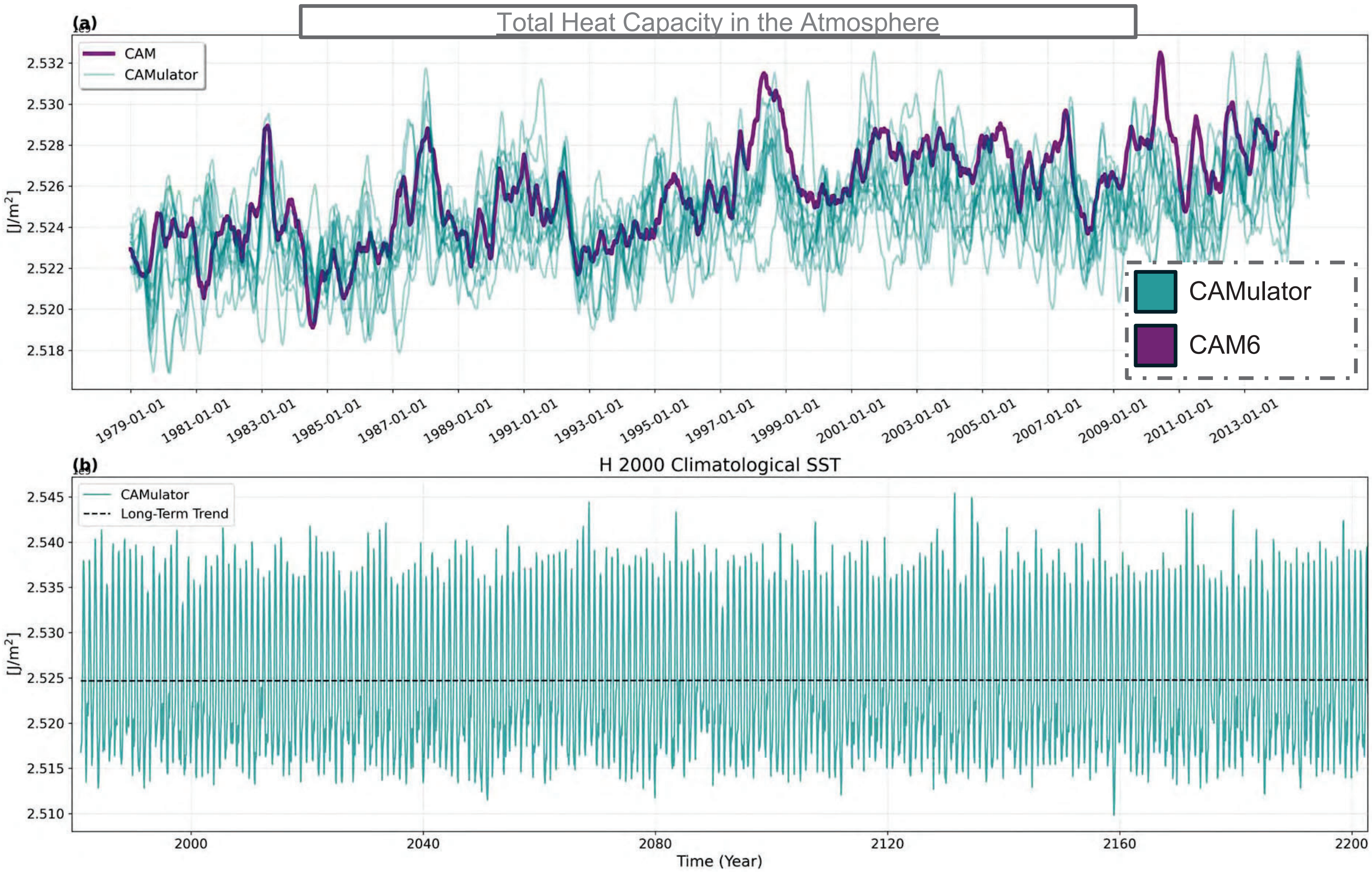
How are the model biases?

CAMulator V.1.0

Response to historical SST forcing

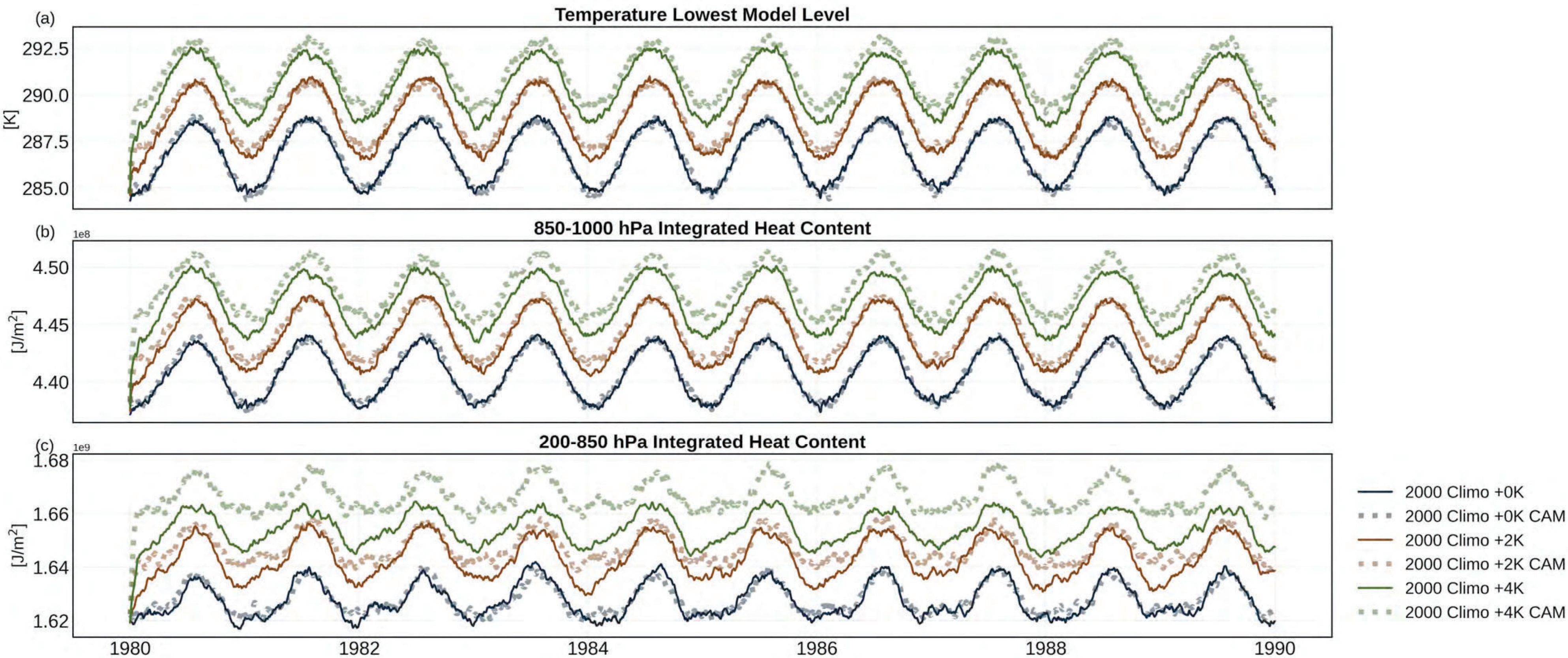
<u>24 Member 35-year ensemble in 8 Hours</u> vs. 4000 Hours for CAM6

Capable of indefinite roll-outs with climatological forcing



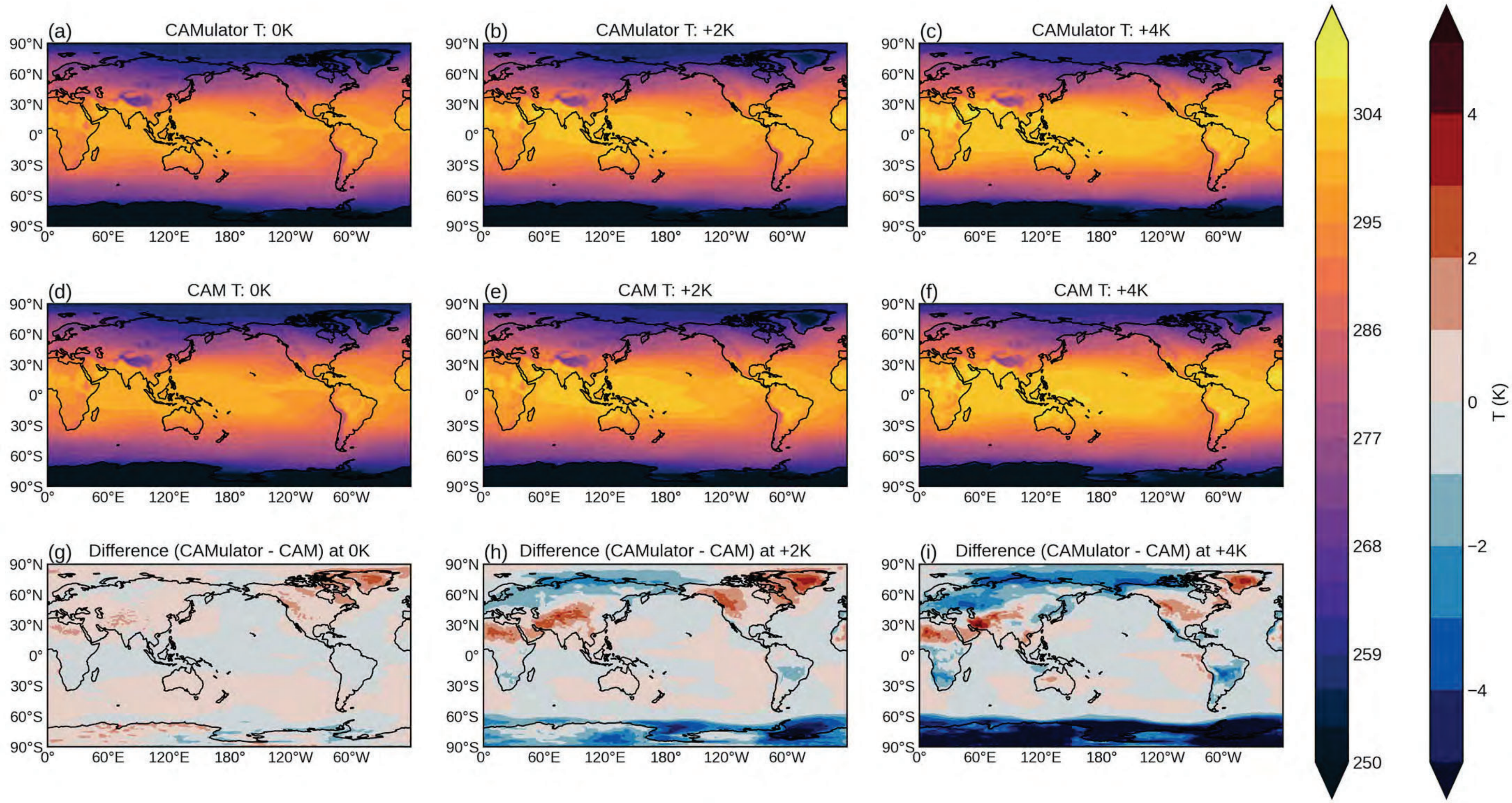
Chapman, Will, et al. "CAMulator: Fast Emulation of the Community Atmosphere … "*arXiv preprint arXiv:*2504.06007 (2025).

Challenge: CAMulatorV1.0 SST Forcing Experiments



Chapman, Will, et al. "CAMulator: Fast Emulation of the Community Atmosphere … "*arXiv preprint arXiv:*2504.06007 (2025).

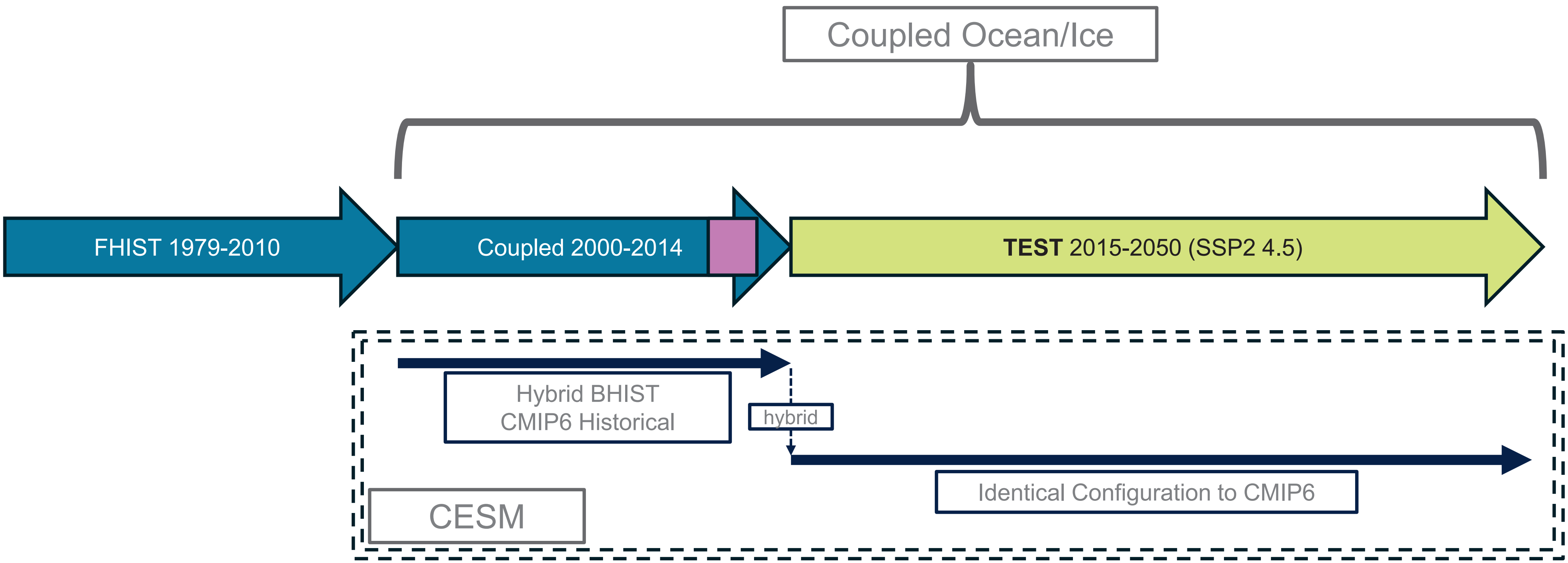
Challenge: CAMulator V1.0 SST Forcing Experiments



Chapman, Will, et al. "CAMulator: Fast Emulation of the Community Atmosphere … "*arXiv preprint arXiv:*2504.06007 (2025).

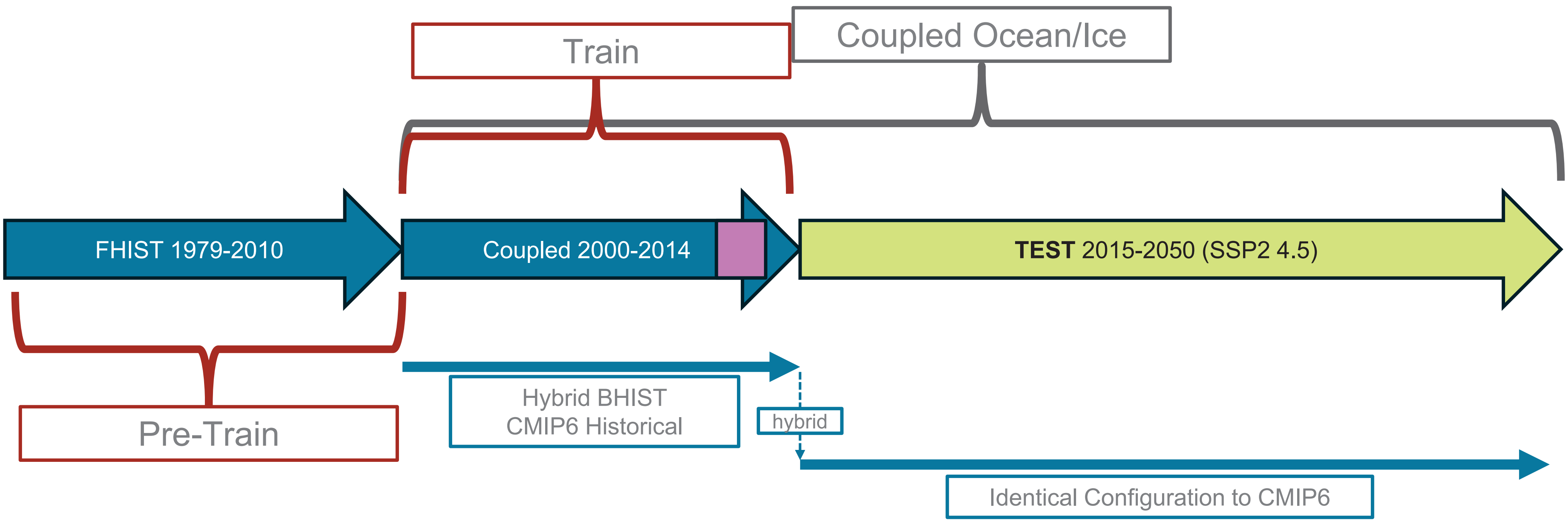
CESM Data + Training Protocol; Updates:

<u>We are now combining FHIST & Coupled Runs in CESM2.1.5:</u>



CESM Data + Training Protocol; Updates:

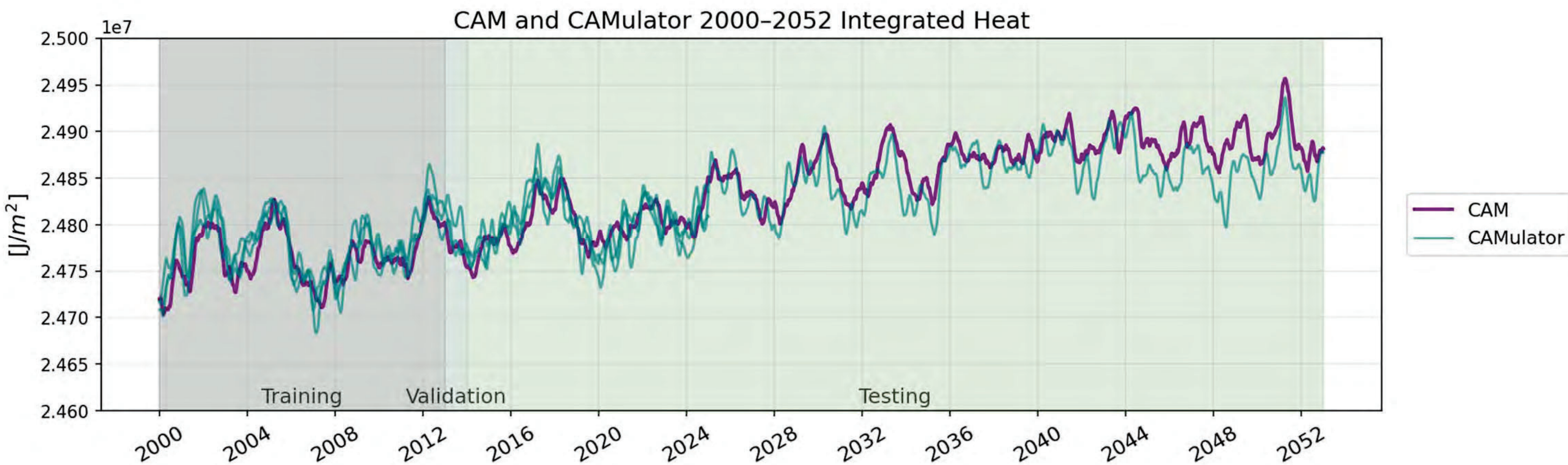
We are now combining FHIST & Coupled Runs in CESM2.1.5:



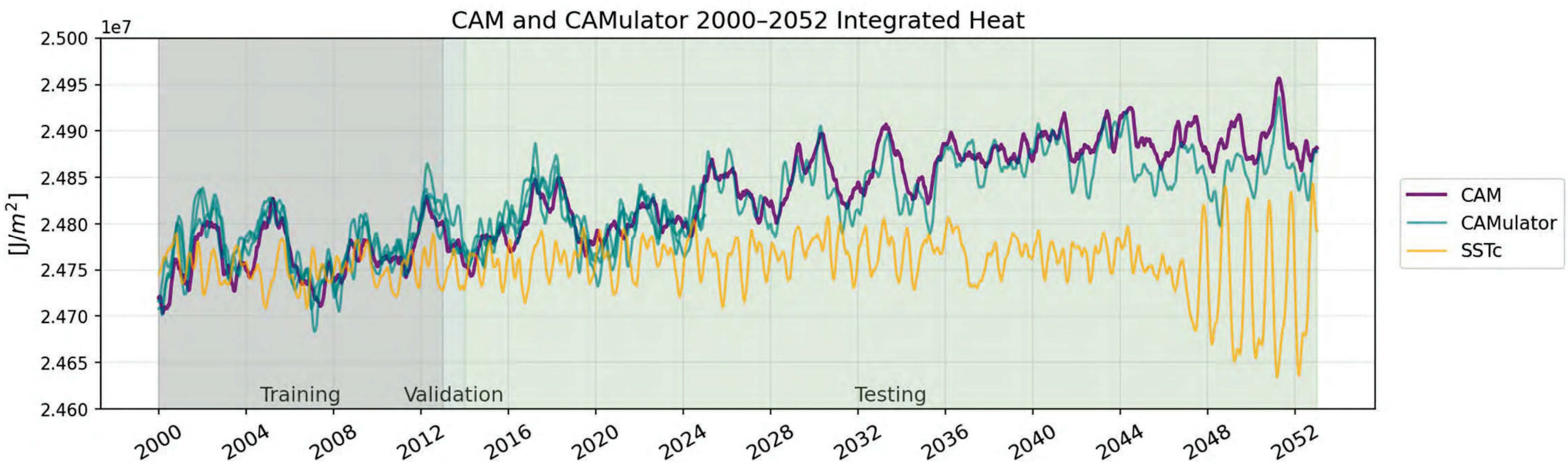
Additional List of Changes from V.1.0:

1. Add Surface Temperature (TS)
2. Add Sea-Ice Fraction
3. Add CO2 Volume Mixing Ratio
4. Add Forcing from Coupled Ocean
5. Remove Redundant Information Variables
6. Remove UpConvTranspose2D Layers

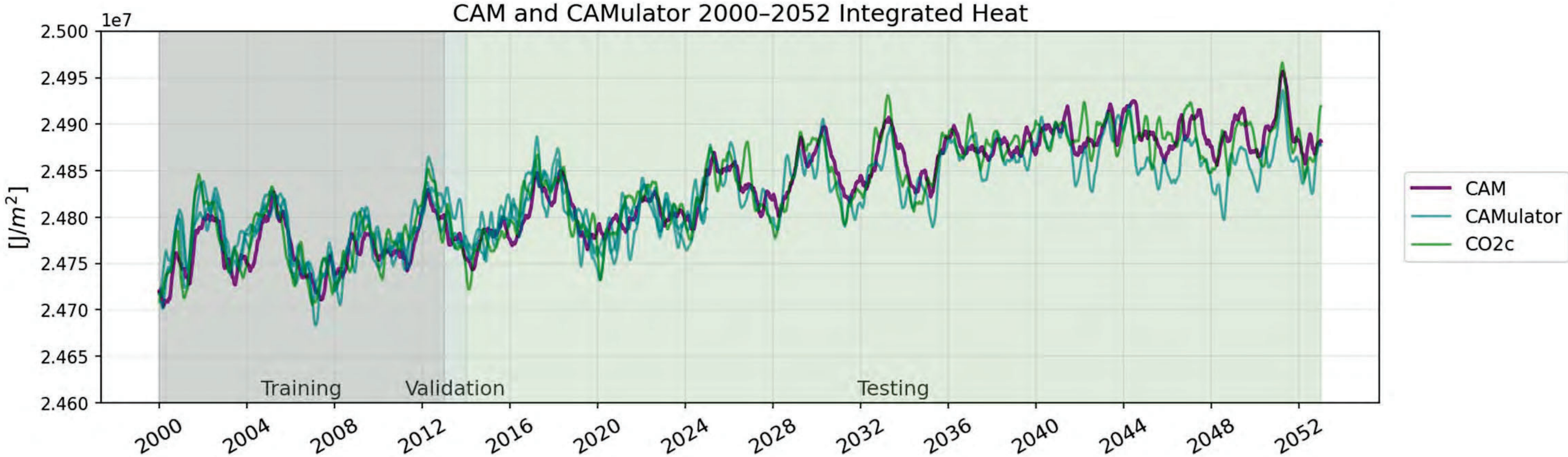
CAMulator V1.1



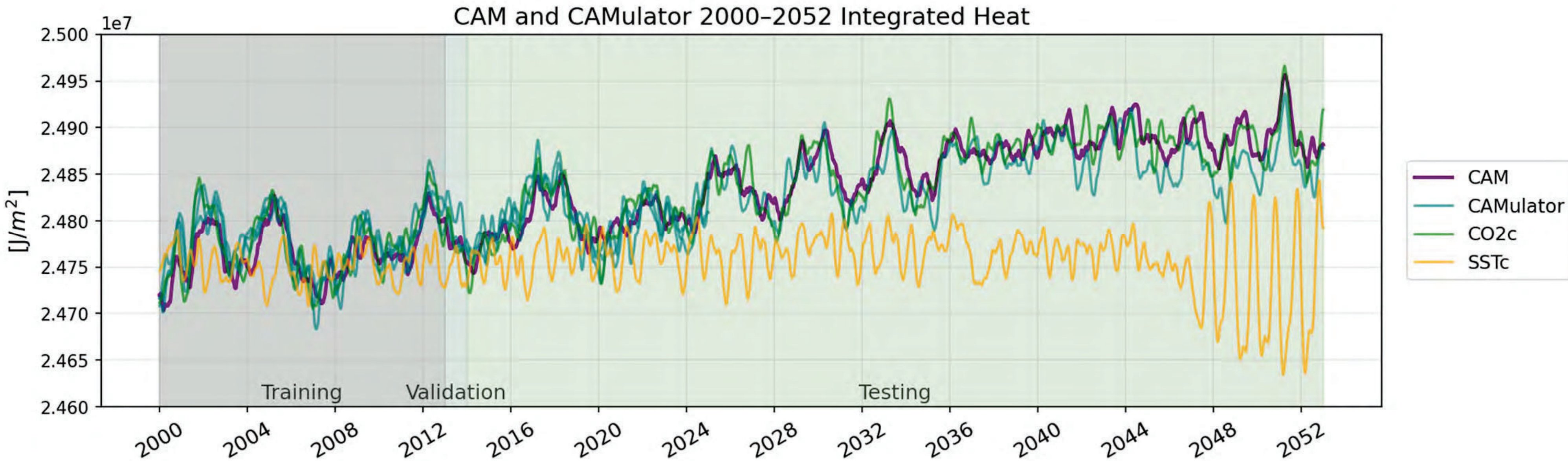
Constant 2010 SST – “observed” CO2



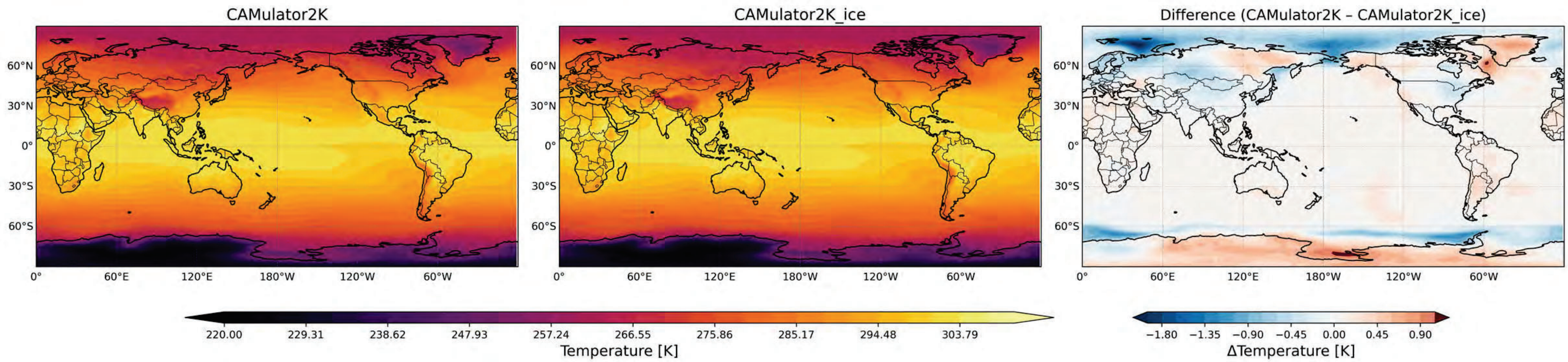
Constant 2010 CO2 – “observed” SSTs



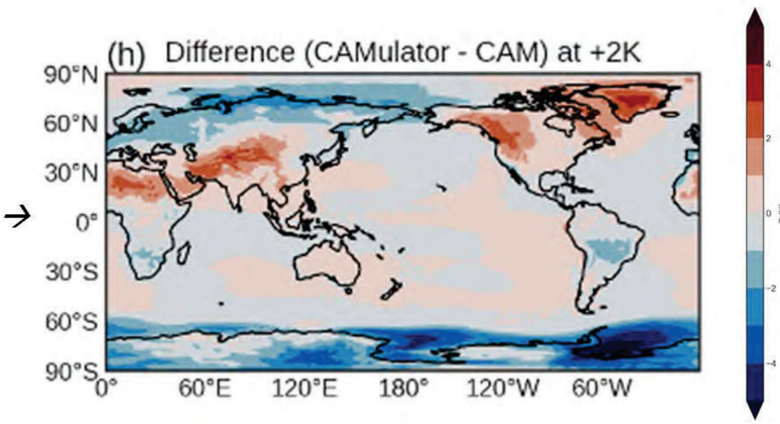
Constant 2010 SST – “observed” CO2



Two Meter Temperature Annual Biases: +2K experiment + Sea ICE



CAMulator V.1.0 Model Biases: →



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?
- Data assimilation: traditional methods versus DA emulators vs hybrid methods

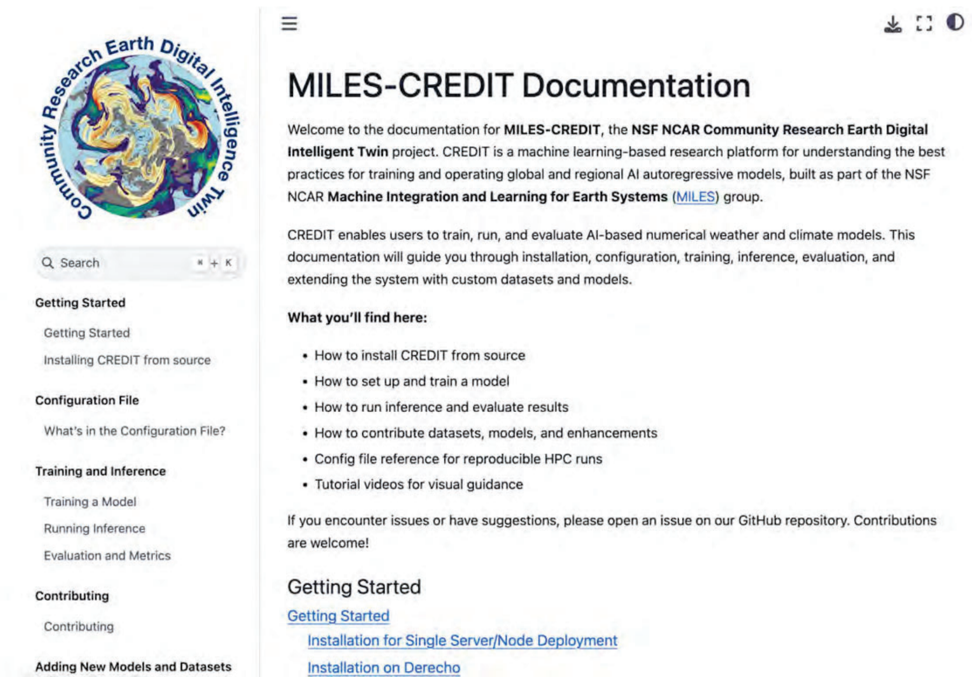
Next Steps

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

Summary

Version 2025.2.0 is out now!

- CREDIT opens a new pathway to customization of the whole AI weather and climate modeling pipeline
- New paradigm of building AI-ready processes and interacting software
- CREDIT source: https://github.com/NCAR/miles-credit
- Links to CREDIT papers: https://miles.ucar.edu/projects/credit/



Contact Me
Email: wchapman@ucar.edu
Github: willychap