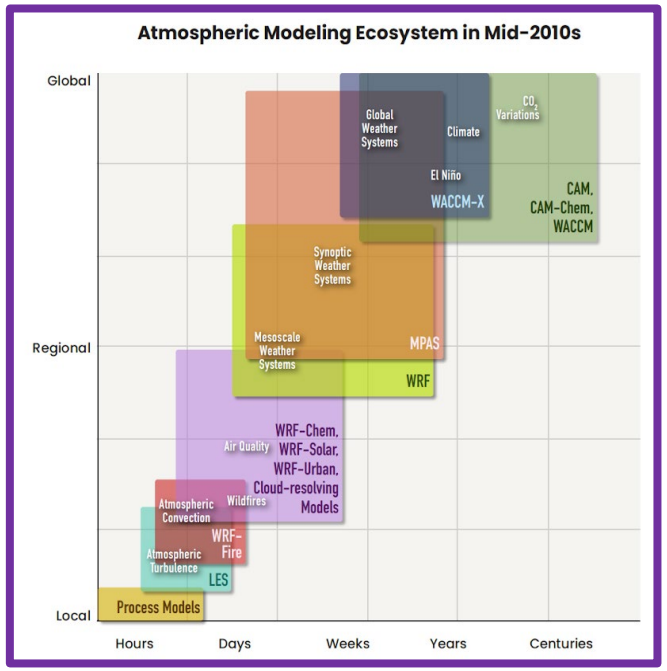
The NSF NCAR Community Software Facility: Transforming Software Engineering for Earth System Models

Thomas Hauser
Director of the Computational and Information Systems Lab
US National Science Foundation National Center for Atmospheric Research (NSF NCAR)

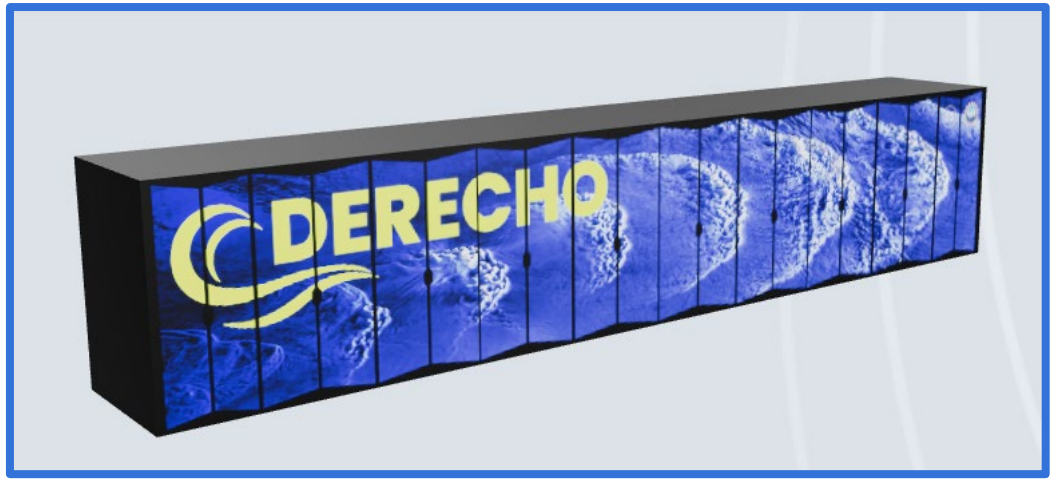
September 18, 2025

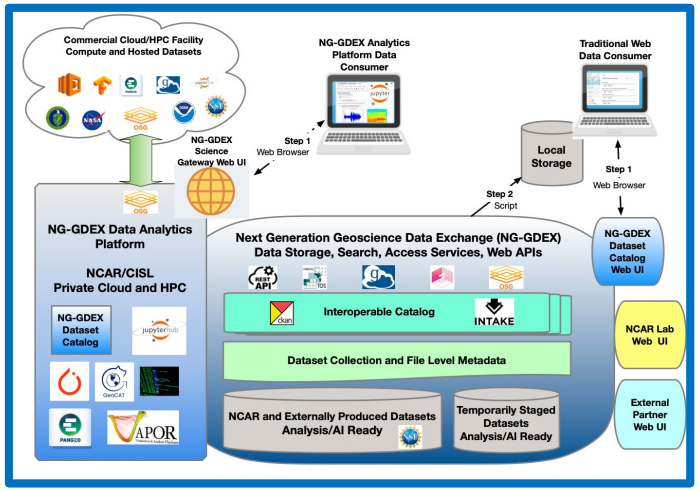
This material is based upon work supported by the National Center for Atmospheric Research, which is a major facility sponsored by the National Science Foundation under Cooperative Agreement No. 1852977.

Research Infrastructure at NSF NCAR



Models and other software as scientific instruments





Cyberinfrastructure – HPC and Data Infrastructure

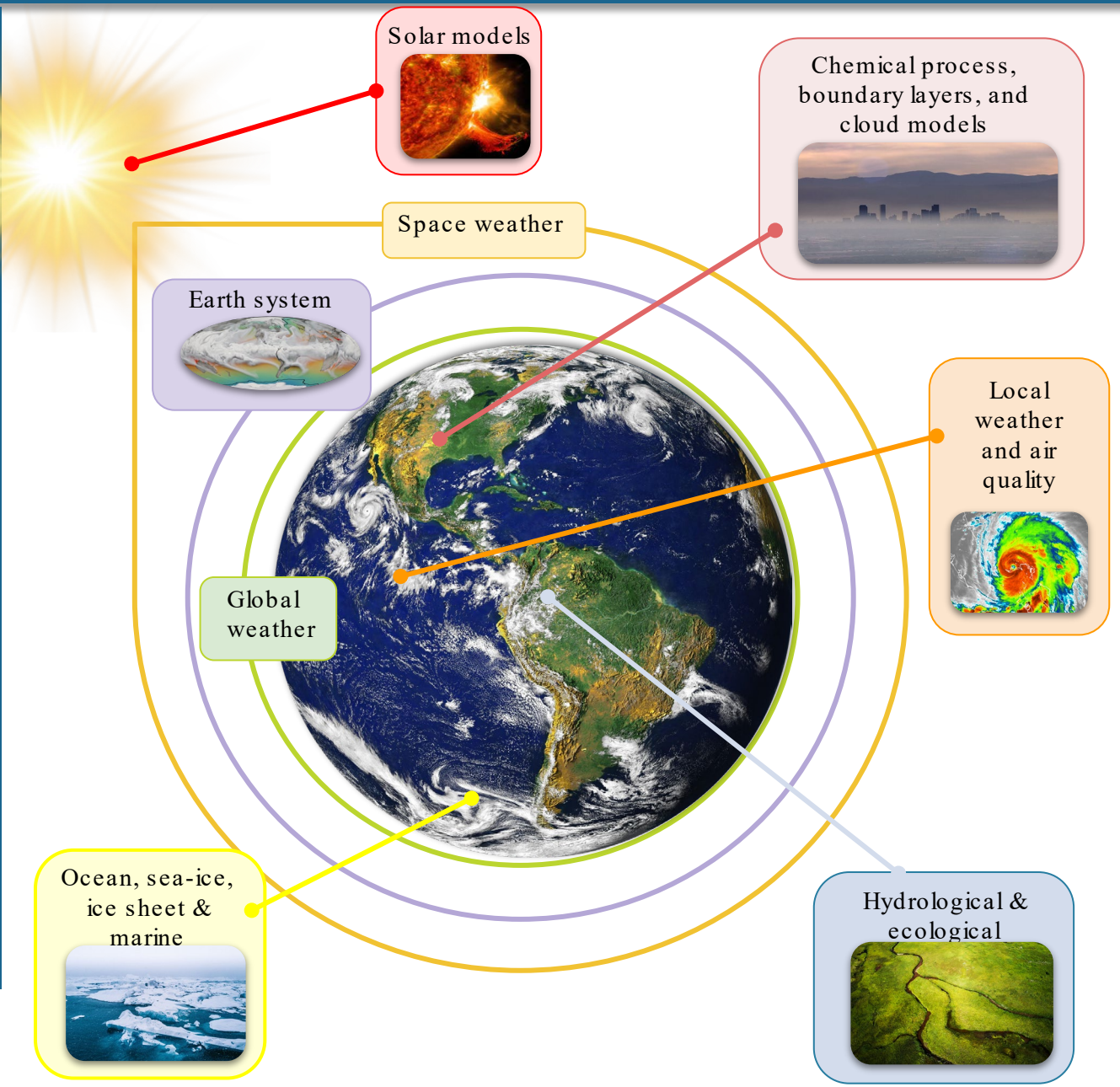
NCAR'S Modeling Ecosystem – Opportunity and Challenge

The collection of models, informed by NCAR's deep scientific expertise, represents a diverse, flexible, and powerful *modeling ecosystem* → enables interdisciplinary Earth system research across.

Most of our software is not ready for modern AI hardware based systems
- No dominant programming model has emerged
- Vendor-specific GPU instruction sets
- Performance portability
- Extensive code refactoring required

Lack of robust commercial compilers (Fortran and C++)
- GCC is best option but no Fortran equivalent

How do support legacy applications?
- WRF, WRF-Chem, …



We are at a breaking point → change is necessary to support the community

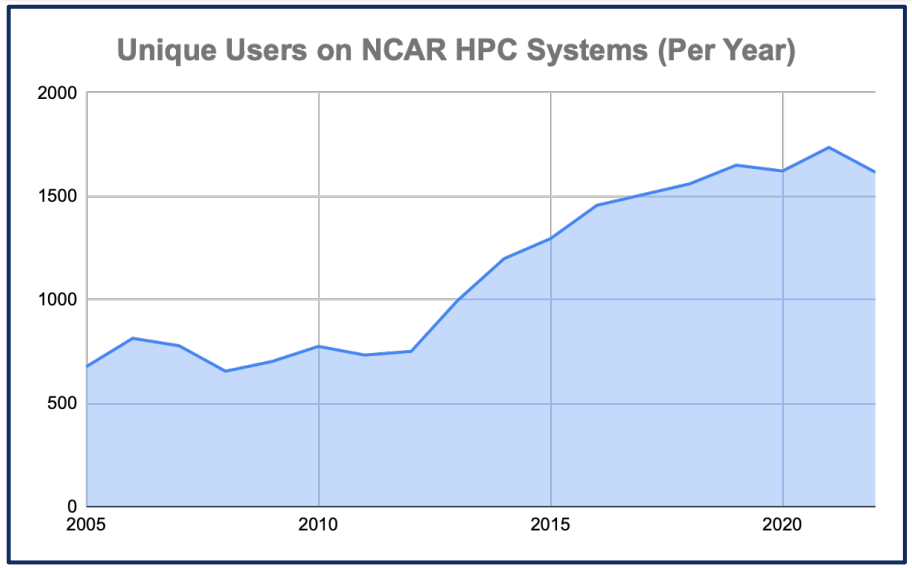
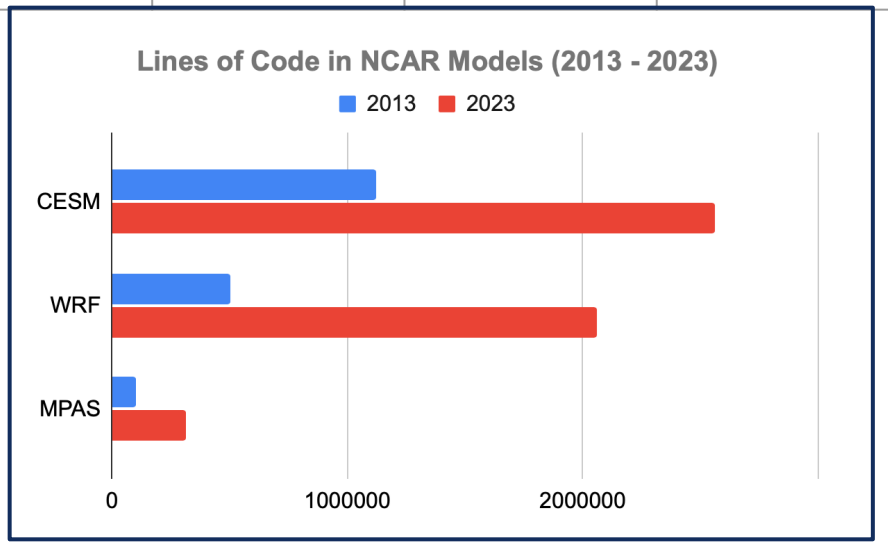


Image from: Nowogrodzki, Anna. 2019. "How to Support Open-Source Software and Stay Sane." *Nature* 571 (7763): 133–34. https://doi.org/10.1038/d41586-019-02046-0.

New Users, Threads and Messages on the CESM Forums (per year)

	2020-2021	2021-2022	2022-2023
Users	420	486	564
Threads	741	869	896
Messages	3609	4182	4630

Model	Open Issues	Closed Issues	Open PRs	Closed PRs
CESM	1190	4003	66	4961
WRF	121	177	10	1513
MPAS	59	127	57	787



CONCEPT: Community Software Facility

1 SCIENCE DRIVER

Earth System Science

2 COMMUNITY DRIVER

Democratization of Science

Reimagining community models as research infrastructure

3 TECHNOLOGY DRIVERS

Compute and Data Technology

Software Development

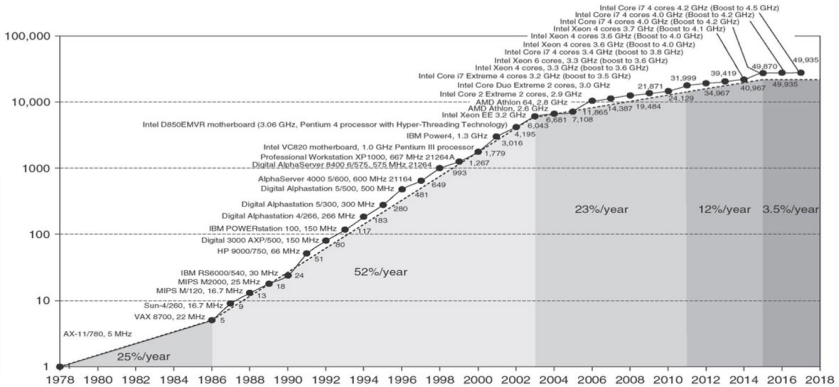
4 WORKFORCE DRIVER

Workforce of the Future

Disruptive Trends

- Societal need
- Increasing model complexity, resolution & ensembles
- Rapid changes in software & technology

Marshall Fire, Superior, CO, Dec 2021



Modernizing NSF NCAR's Model Ecosystem – Community Software Facility

Vision for the Community Software Facility

- Manage best practices and governance of our "core" community software
- Support the use of our software across leadership class hardware systems
- Resourced to maintain and evolve our software while avoiding toxic technical debt

Project to create the facility "Discovery project for the CSF"

- Develop NCAR-wide software governance
- Change the current culture and create a community of practice
- Provide supported tools and standards
- Deliver new modeling capabilities

It's a culture change activity

The Community Software Facility Leadership Team

CSF Governance

Thomas Hauser
CISL Lab Director, CSF Executive Sponsor

Gretchen Mullendore
MMM Lab Director

Glen Romine
Senior Advisor for Research and Strategic Planning

Jon Petch
CGD Lab Director

Program Management

Sheri Voelz
CSF Discovery Project Lead & CISL Applied Computational Science Section Manager

Domi Colegrove
Strategic Initiative Projects Lead, FLPO

Kate Kuzemka
Project Manager, FLPO

Community Software Facility Discovery Project

Understand priorities and challenges for the community software facility

- SEAL team recommendations
- Discovery Workshop: Scientific software best practices, tools, and culture – Aug 21-22

Governance

- Decision making processes
- Internal coordination
- External collaboration

Culture & Community of Practice

- Culture improvements
- Empowering SEs
- Training and professional development
- Transform the Software Engineering Assembly into a Community of Practice

Software Engineering Tools & Approaches

- Development standards and best practices
- Software productivity
- Interoperability
- Computational performance

Capabilities

- Atmosphere/ocean coupling
- Interoperable physics and dynamics
- Interoperable chemistry

Software Engineering Tools and Approaches

Performance and GPU programming models

- Separation of concerns
- Kokkos
- AMReX
- OpenACC
- Python, Dace, GT4py
- I/O – compression, asynchronous

Supported CI/CD environment

- Enabled by CIRRUS: Cloud Infrastructure for Remote Research, Universities, and Scientists
- Sets up the CI workflow through GitHub Actions
- User-defined container image to Harbor container registry
- Creates a GitHub personal access token
- Sets up the secrets management (OpenBao)

Software Architecture and Software Engineering Productivity

- Create an NCAR-wide software architecture team
- Evaluate AI for code optimization and transformation

Testing

- Unit test
- Statistical testing

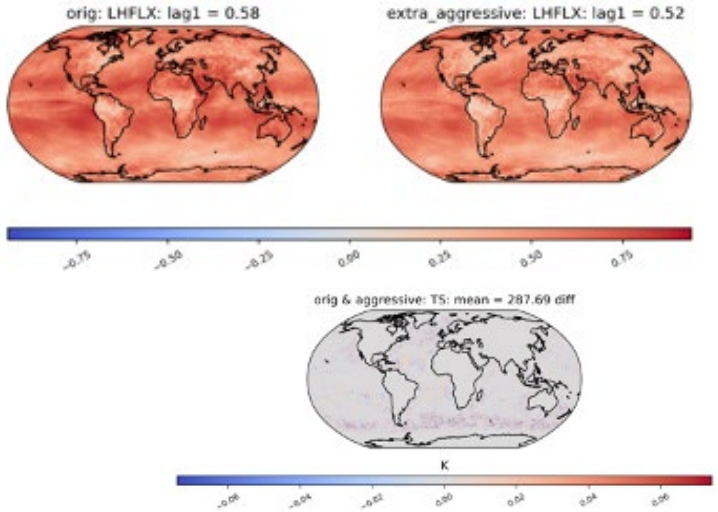
Lossy compression for Earth system models

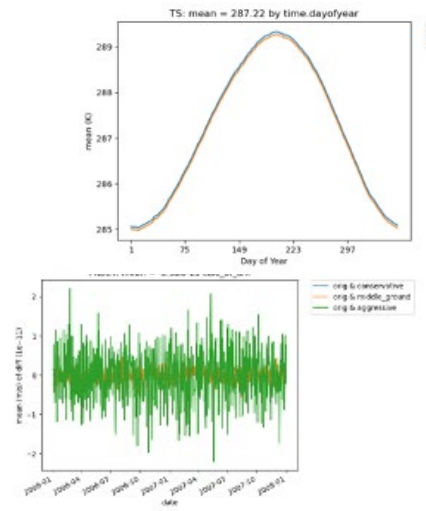
Data generation trends at NCAR are unsustainable!

INCREASING COMPUTE POWER → MORE AND MORE DATA!

Storage resources are limiting climate science objectives!

Goal: Use lossy compression to reduce CESM storage …without (negatively) impacting science results!





Challenges for climate data:

- Spatial & temporal dependencies
- Output variables have very different characteristics
- Typical compression metrics don't have much meaning
- Parallel I/O

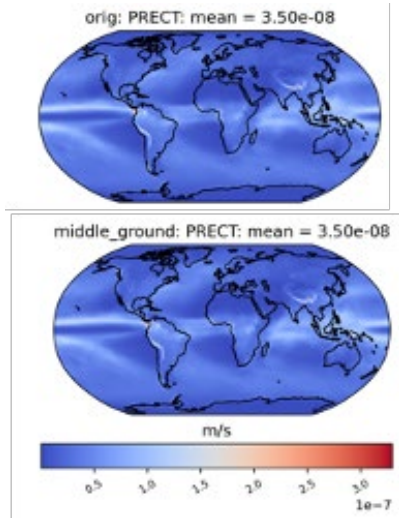
Research focus areas:

(1) How to evaluate the effects of lossy compression on CESM data?

- Quantifiable metrics (and thresholds) for derived quantities and statistics
- Python package (LDCPy) for analysis of differences in large spatio-temporal datasets

(1) From features of the data, determine what type of (and how much) compression to use?

- Automated tool to determine the optimal algorithm and compression rate for each variable!



Ensemble Consistency Test (ECT)

Quality assurance: Changes during the Model development cycle should not adversely affect the results!

But many changes are not Bit-for-Bit (e.g., compiler change, optimization, new hardware) …

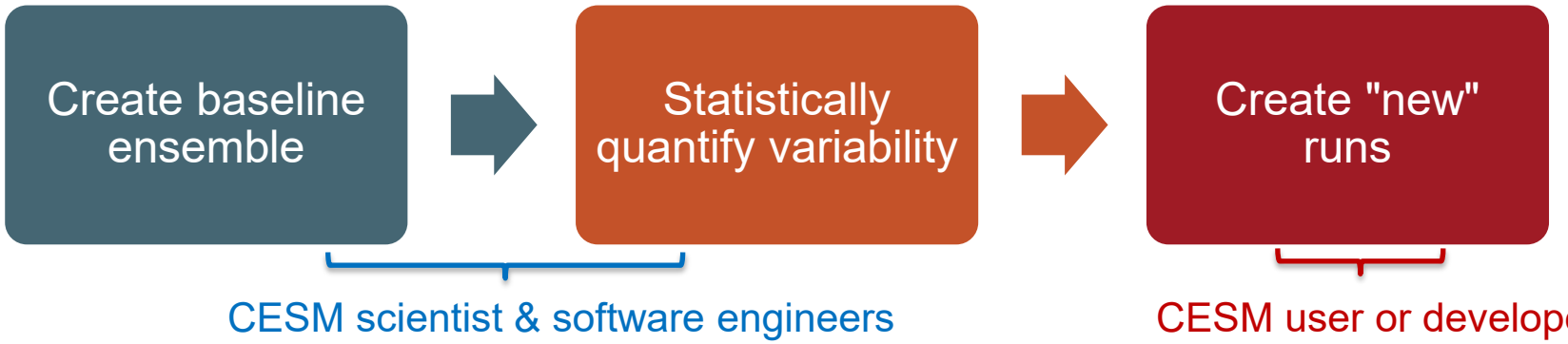
Question: Is the new answer *correct*? (not well-defined)

Alternative question: Is it *statistically distinguishable* from the original?

ECT approach: evaluate in the context of the climate model's variability

(climate scientists think of uncertainty through the use of ensembles)

https://github.com/NCAR/PyCECT



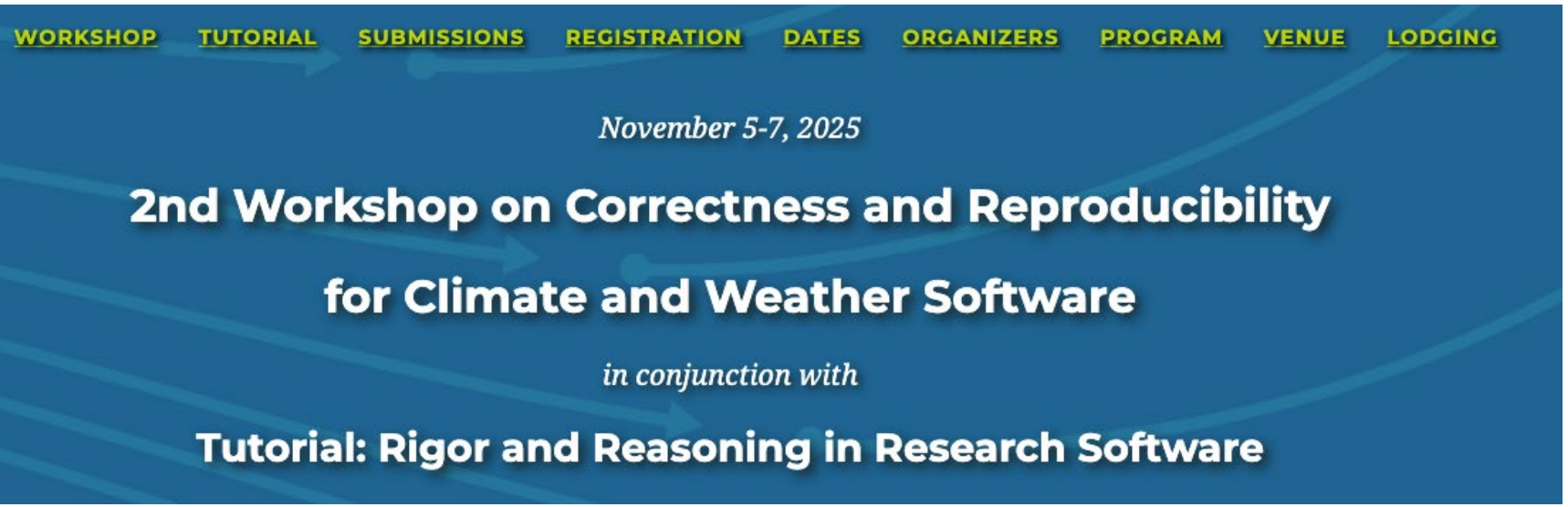
Works extremely well in practice!

- Used by CESM for ~ 8 years & now also MPAS
- Modifications *expected* to be climate-changing *fail*
  - e.g. relative humidity, dust emissions, CO_2 levels
- Modifications *not expected* to be climate changing *pass*
  - e.g., threads, -O0, compiler version, code rearrangement

Future Work

- Why did the ECT fail? (i.e., what caused the statistical difference?)
- Adapt ECT approach to MOM6
- Explore correctness approaches for individual physics parameterizations

Questions



- Credits
  - Alison Baker
    - Compression & ECT
  - Jon Petch, Gretchen Mullendore, Glen Romine
  - Domi Colegrove, Kate Kuzemka

https://ncar.github.io/correctness-workshop/

Registration (in-person or virtual) is open and program is available!