CESM Software Engineering Working Group

A selection of major activities

Bill Sacks
CESM SEWG co-chair
Credits: Many members of the SEWG and others

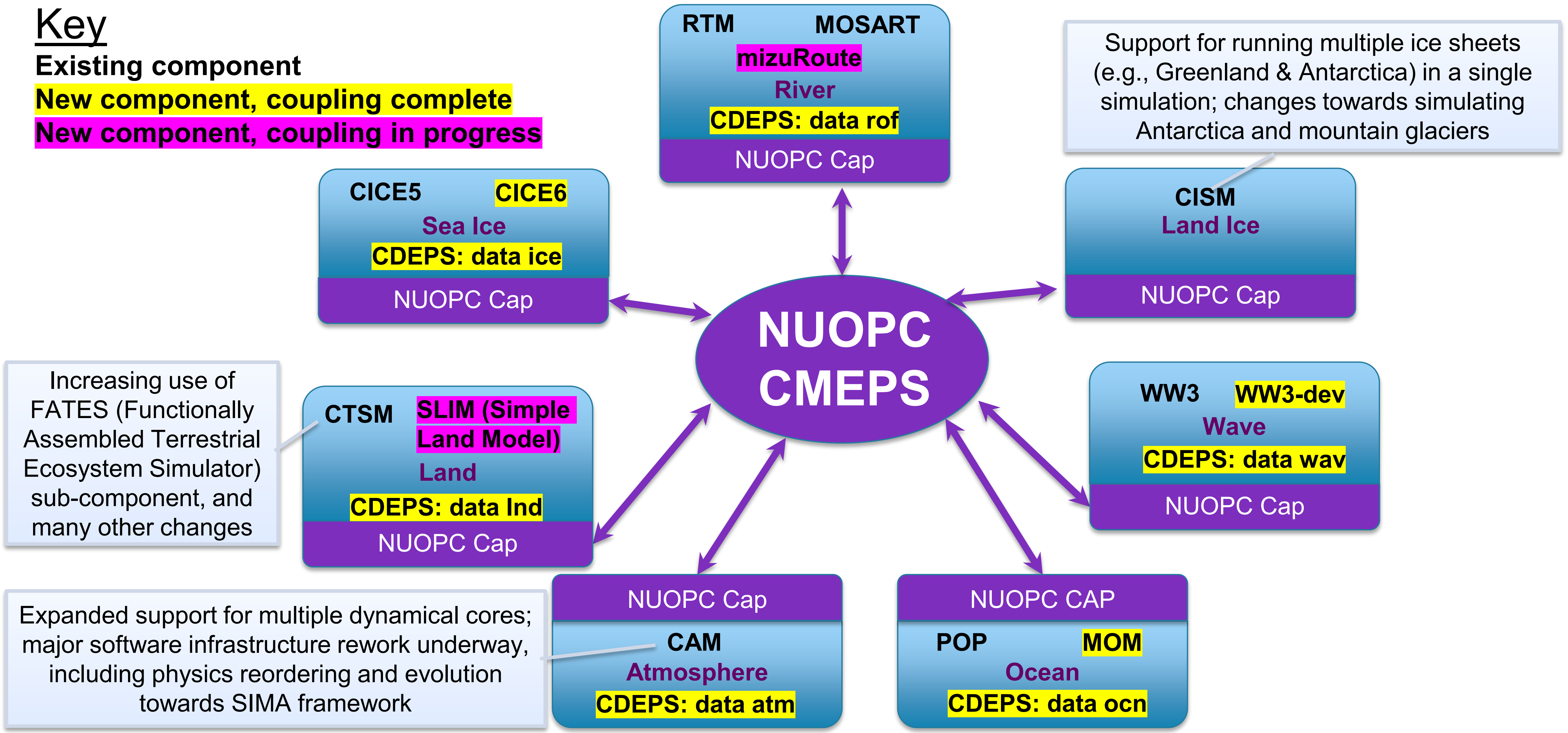
NCAR

13 JUNE 2022

NSF

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.

New / Significantly Expanded Components



New Infrastructure Repositories

- CIME (https://github.com/esmci/cime/) is now mainly limited to the python-based Case Control System (jointly developed and used by CESM & E3SM)
- Fortran-based model infrastructure now resides in its own repositories:
  - CMEPS
  - CDEPS
  - CESM_share
  - CPL7 (deprecated)
- Configuration data is now in CCS_config
- For pointers to each infrastructure and component repository:
  - https://github.com/ESCOMP/CESM/blob/master/Externals.cfg
  - Or click on "New issue" from https://github.com/ESCOMP/CESM/issues to get links to each repository's issues page

CMEPS / CDEPS Updates

(more details Wed. at 8:35 am)

- **CMEPS / CDEPS is now the default coupling infrastructure & data models for CESM**
  - Extensive validation was done, including multi-century simulations
  - The MCT-based infrastructure will soon be deprecated in development code; still supported in CESM2.1 (CMIP6) release code

- Updates to CMEPS coupling:
  - Atmosphere-ocean flux calculation can now be done on the exchange grid
  - Added exchange of enthalpy fluxes – needed for energy conservation with MOM
  - Multiple ice sheets (e.g., Greenland and Antarctica) can be run simultaneously, controlled at runtime
  - Can couple multiple ocean layers with ice sheets

- CDEPS data model functionality now extensively used in multiple components (e.g., nitrogen deposition read by CAM)

CMEPS / CDEPS Benefits

- Jointly developed by NCAR and NOAA
  - Has enabled sharing code and caps for CICE6, MOM6 and WW3 with NOAA
  - Testing the coupling and data models in multiple modeling systems leads to a much more robust code base
  - CDEPS's use by NOAA is bringing in support for new forcing data streams
- Much easier to introduce new grids
  - Most mapping generation now done at runtime: **now only need 4 pre-generated mapping files** (for custom runoff mapping) **instead of 25**
  - No longer need domain files to specify land fraction
- Can transfer 3-d fields and related fields as a single packaged field
- CDEPS provides ability to do 3-d mapping from input streams and provides many new mapping types

ESMF Updates

A selection of enhancements in the latest ESMF releases:

- All regridding methods are now supported on **exchange grids**
- Progress towards adopting the Mesh-Oriented datABase (**MOAB**) library for internal mesh representation
- Update "creep fill" **extrapolation** method
- Flexible **NUOPC alarm specification**, allowing model phases like writing of restart files to be executed less frequently
- New options for detailed **performance profiling** of NUOPC components
- **Upgrade to ParallelIO** (PIO) from v1 to v2
- Various **performance optimizations**: Per-component threading levels, scalable mesh creation from file, and others

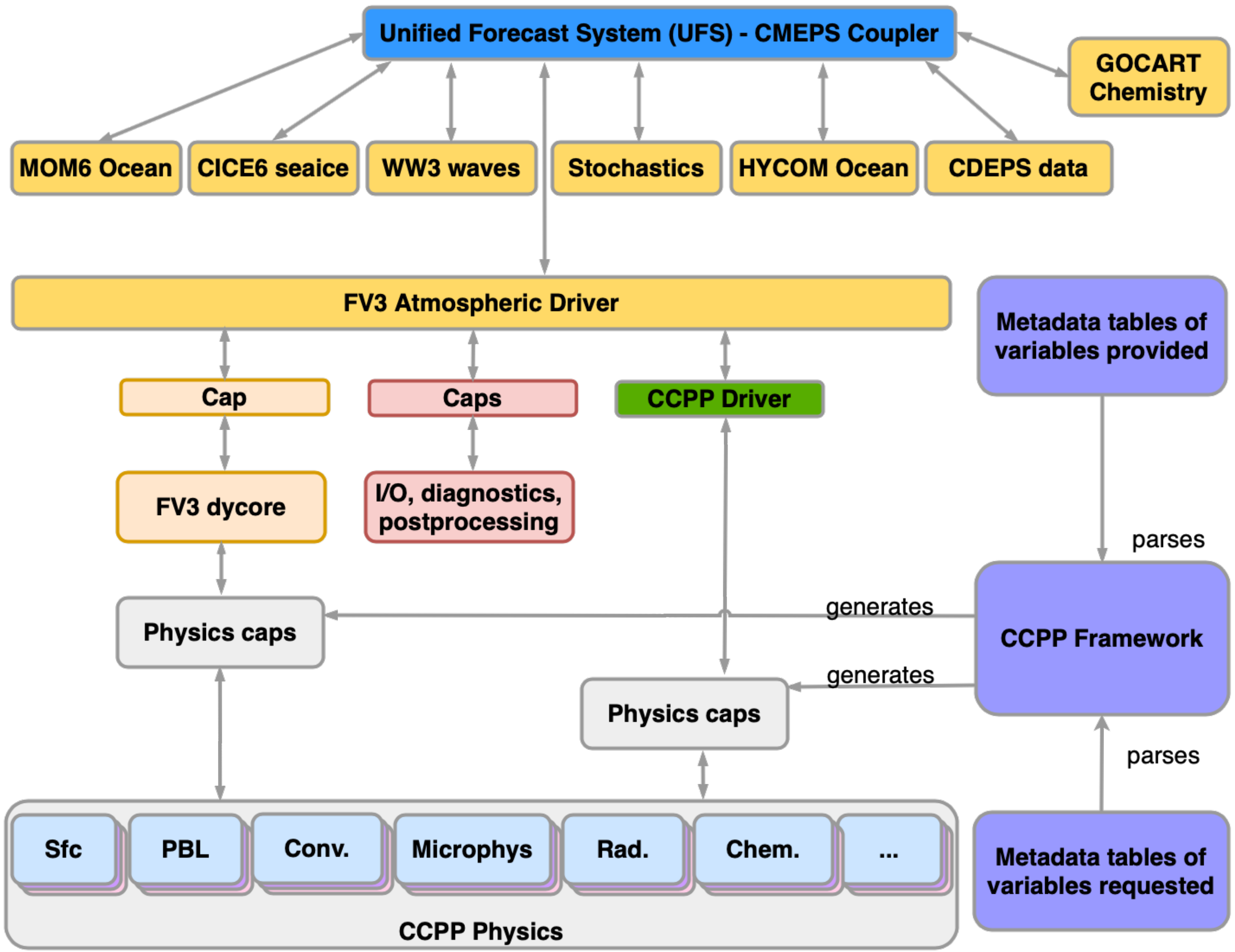
Improved Threading Control in CESM

- Previously, if OpenMP-based threading was used, all components needed to use the same number of threads
- Now each component’s thread count can be controlled separately
- Allows for finer-grained control of load balancing and improved scalability: some components perform better with threading, whereas others perform better without threading
- Enable by setting the xml variable `ESMF_AWARE_THREADING=TRUE`

Ultra-high Resolution CESM

- Targeting a 3.75km global CAM configuration
  - Have successfully run a 7.5km F2000 (coupled atmosphere-land with data ocean) case; needed to resolve several memory bottlenecks
  - Currently working through additional challenges exposed with 3.75km
- Also working on a similar configuration coupled to MOM at high resolution with the Texas Advanced Computing Center (TACC)

Common Community Physics Package (CCPP)



(more details Wed. at 9:05 am)

- CCPP contains a **library of physical parameterizations** and a framework that connects it to host models
- It is used by various **host models**: the CCPP Single-Column Model, the Navy experimental NEPTUNE model, and by the Unified Forecast System. It is on track for transition to **NOAA operations** in 2023.
- **MPAS/WRF** now have a suite of CCPP-compliant parameterizations that can be executed directly
- CCPP Framework is under **further development by NCAR/CGD** to meet additional NCAR requirements

Clouds, Containers and Training

(more details Wed. at 9:25 am)

Cloud Updates:
- Secure multi-user JupyterHub deployment on-demand
- 'EASE' Kernel (preinstalled conda environment)
- CESM ready to run on AWS HPC6a instances
- Persistent accounts via email addresses: will help enable community-based training

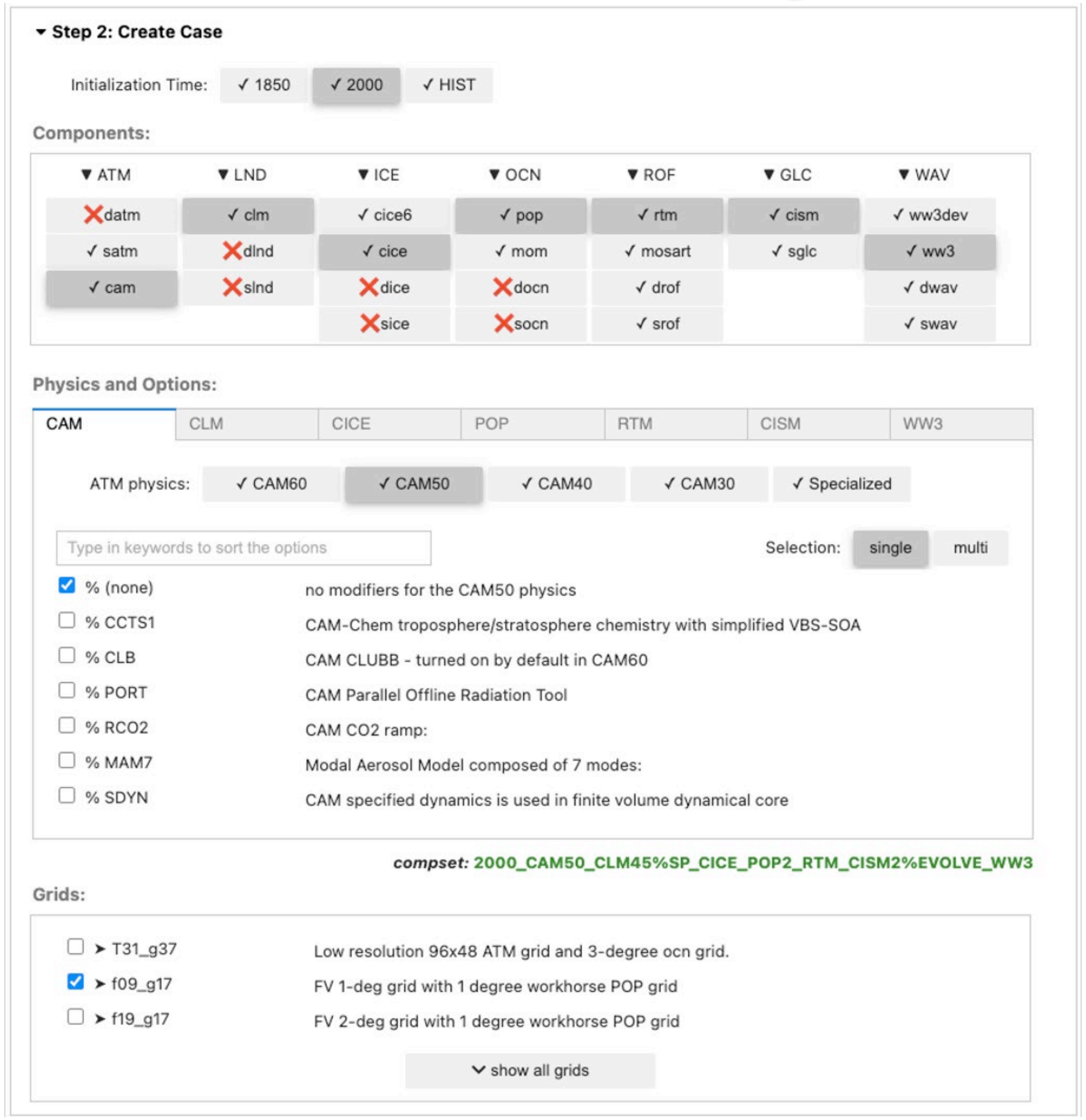
Container Updates:
- CTSM-Lab / CESM-2.3-based containers
- Arm M1 versions
- Updated to use EASE kernel (soon)

GUI & Tools to Support CESM Simpler Models & Custom Configurations

- Graphical user interface guides users through the process of creating CESM cases: choosing appropriate compsets & grids

- New metadata and logic module to check compatibility of compset options and grid

- New land model tools to facilitate creating surface data sets for custom grids and configurations, including idealized configurations

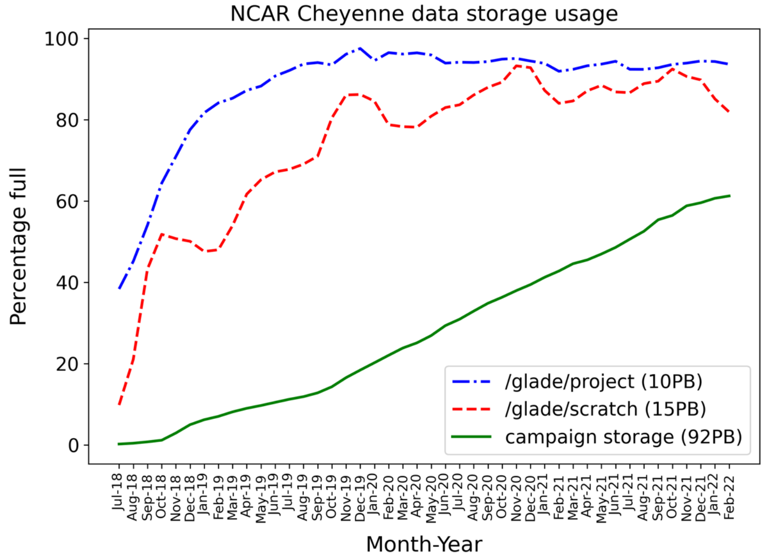


Primary SEs: A. Altuntas, S. Levis
Funded by an NSF CSSI award. (PIs: Bachman, Simpson)

Lossy Compression and CESM Data

(more details Wed. at 10:45 am)

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

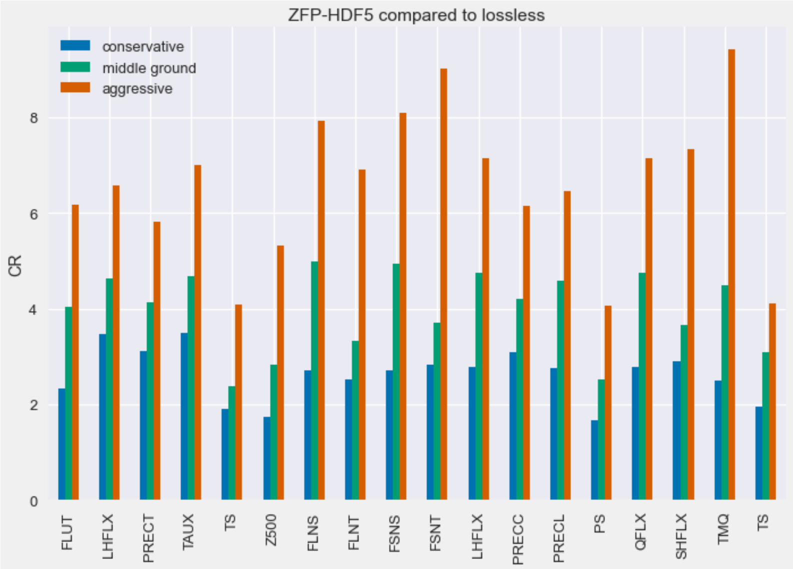


Challenges:

- Compression affects fine spatial and temporal scales
- CESM data diversity: “one-size-fits-all” approach not optimal

Our focus:

Evaluating the effect of lossy compression on CESM data via spatio-temporal statistical analysis tools that emulate the key aspects of climate data analysis (e.g., LDCPy) in order to predict optimal compression



Current work:

- DSSIM (Data Structural Similarity Index Measure)
  - Newly developed to apply directly to floating-point data
  - Indicates whether images generated from the data are likely to have noticeable differences
- A tool for auto-selection of compressor and parameters
  - Using features of the data, can we say something about what type of (and how much) compression to use?
  - We compare statistical models using explicit features and deep learning approaches

A. Baker, A. Pinard, D. Hammerling and H. Xu

New Diagnostics Packages

New Diagnostics Packages

- Atmosphere Diagnostics Framework (ADF): New CAM diagnostics package (https:// github.com /NCAR/ADF)
  - The ADF can now replicate most core features of the old AMWG diagnostics (some plot types and observational datasets are still under development) (more details Wed. at 10:25 am)
  - The ADF will be required for CAM7 / CESM3: it is vertical -level agnostic so can manage the new CAM7 vertical levels
- Climate Variability Diagnostics Package (CVDP)
  - CVDP v5.2.0 was released in Fall 2021, and has been wrapped into the ADF
  - The CVDP can now read unstructured grids from atmospheric/land components
- Ocean Model Diagnostics Package
  - Interim diagnostics package in place for MOM development
  - So far only used internally; more work needed to make it accessible to the community
- Land Ice Diagnostics Package
  - New diagnostics package in early development and planning phase
  - External collaborations through the Ghub.org Glacier science hub and the Ice Sheet MIP (ISMIP) will provide avenues for the community to benefit from this work