DART and CESM: Current Status

Nancy Collins
NCAR Data Assimilation Research Section

With thanks to the DART Group (Jeff Anderson, Tim Hoar, Kevin Raeder) and the CESM Software Development Group (Mariana Vertenstein, Tony Craig, Jim Edwards)
What is Data Assimilation?

Observations combined with a Model forecast...

Ensemble DA produces a set of ‘equally-likely’ analyses.

...to produce an analysis (best possible estimate).
Ensemble Data Assimilation

• Estimate observation values from N model states, compare to actual observation
• Difference of ensemble of estimated observations and actual observation is used to update ensemble of states
• Ensemble sizes of 20-100 members to accurately sample distributions
New Since Last Year

• CESM supports multiple **active** Multi-instance components
• With CESM 1.1 Beta15 release **all** components have consistent namelist/configuration support
• CESM short-term archiver updated to handle N instances of output files
• POP
  – First CESM component to support Multi-instance capability; running various ongoing assimilation projects
• CAM
  – Routinely running CAM assimilations with CESM Multi-instance
• CLM
  – Doing initial assimilations with CESM Multi-instance
CAM Assimilation with CESM Multi-instance
DART

Ocn Obs

3D state

3D restart

POP

DATM

2D forcing from CAM assimilation

2D forcing

Coupler

POP Assimilation with CESM Multi-instance

CESM SEWG; 19 June 2012
CLM Assimilation with CESM Multi-instance

Coupler

DATM

2D forcing from CAM assimilation

2D forcing

CLM

3D restart

3D state

DART

Lnd Obs

3D forcing from CAM assimilation

NCAR

CESM SEWG; 19 June 2012
Next Steps:
Fully-coupled CESM with DART assimilating observations into a single component
Next Steps:
Multiple components assimilating with different DART(s) in fully-coupled CESM
Next Steps: Fully-coupled CESM with full-cross-component DART assimilation

All types of Obs
Longer term goal is to avoid file I/O when passing data between CESM and DART.

This will require the coupler to pass subdomains of 3d fields (N instances of them!) instead of the 2d fields it passes now.

The decomposition won’t just be inside a single grid, but will be across the N instances of a 3d subdomain.
Challenges and Opportunity

• Yellowstone (and other large clusters) has less memory per node but more nodes
  — Communication costs are high if all tasks involved
  — Using intermediate files will kill performance
  — Forward operator computation on partial states

• DART group has 1-year term position open now for either SE II/III or PS I/II to work on the next generation of many aspects of the DART software package
Web and Contact Info

- [http://www.image.ucar.edu/DARes/DART](http://www.image.ucar.edu/DARes/DART)
- General questions: dart@ucar.edu
- Or contact the DART team members directly:
  - Jeff Anderson: jla@ucar.edu
  - Kevin Raeder: raeder@ucar.edu
  - Tim Hoar: thoar@ucar.edu
  - Nancy Collins: nancy@ucar.edu