Data Assimilation with Multi-Instance CESM Components and DART

K. Raeder, J. Anderson, T. Hoar, N. Collins, A. Karspeck, A. Fox

National Center for Atmospheric Research
Boulder, Colorado, USA
dart@ucar.edu

1. DART and Multi-Instance CESM

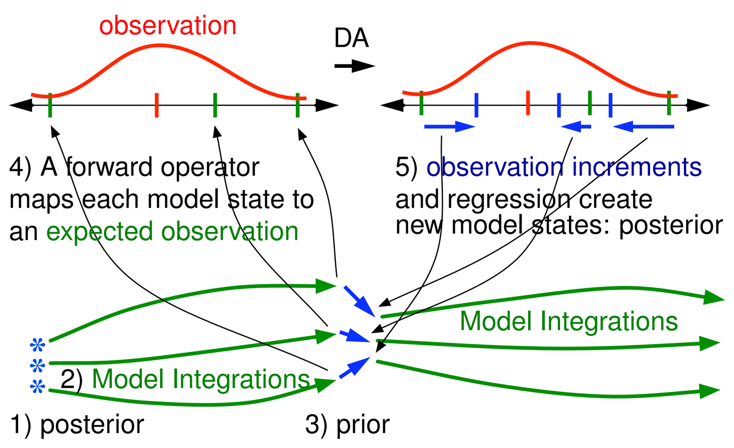


Figure 1: *Illustration for a toy ensemble size of 3.*

CESM can advance multiple instances of one or more model components simultaneously, which enables it to use DART to assimilate observations and bring the model state(s) closer to the truth.

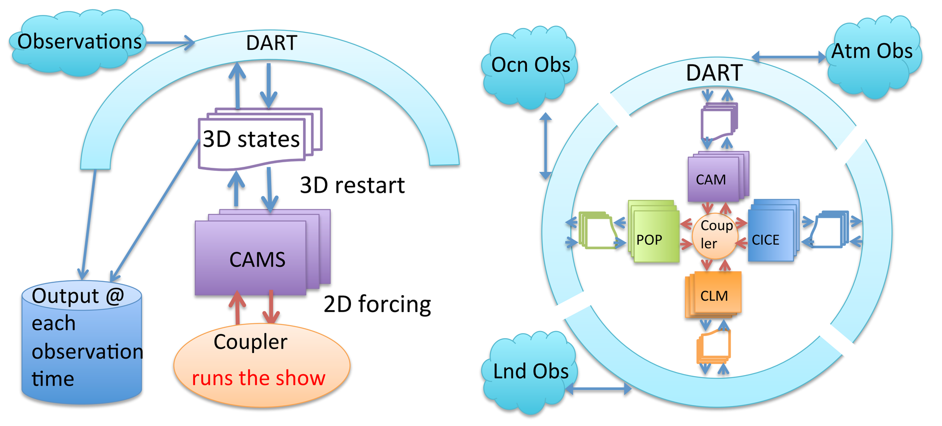


Figure 2: *In the new paradigm, the coupler runs the experiment, advancing CAM when necessary, while DART runs when there are observations to assimilate.*

2. CLM Single Column Preliminaries

The interface between CLM and DART has enabled assimilation of MODIS snow cover observations into CLM, and investigations are underway to assimilate flux tower data into CLM run in single-column mode. A tower provides only 1 representation of atmospheric forcing, but ensemble data assimilation needs an ensemble of forcings. A natural source of this is an ensemble of CAMs, but the following question must be answered. Can CAM provide meteorology which is close enough to the tower meteorology?

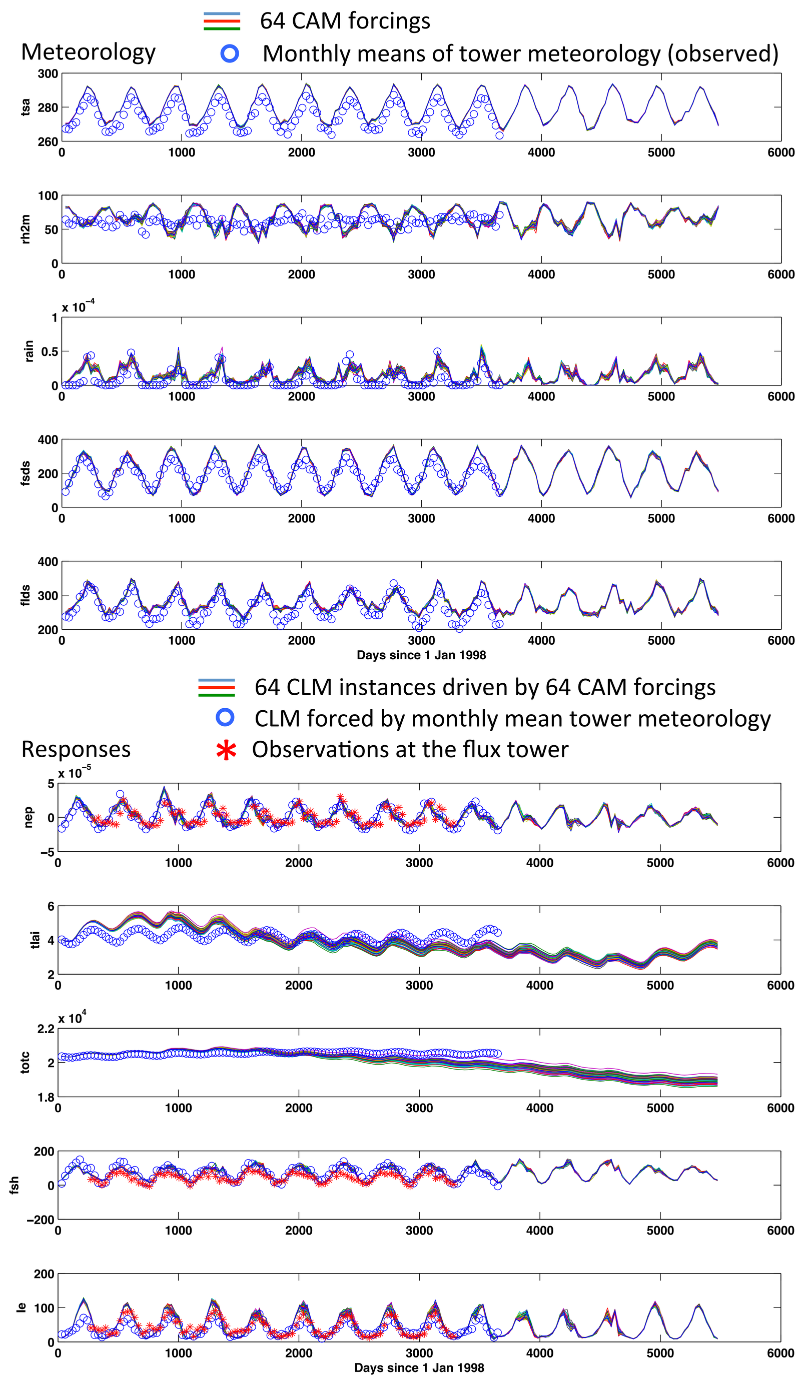


Figure 3: *Comparison of free-running CLMs forced by meteorology from the Niwot Ridge flux tower (blue circles) and by an ensemble of 64 CAM columns at a nearby model gridpoint (colored lines). No observations were assimilated.*

3. Global Atmosphere Assimilation

DART assimilated all the observations that were used in the NCEP/NCAR Reanalysis into CAM4 to produce a global, 6-hourly, 80-member ensemble reanalysis for 1998 through 2010 — with plans to continue. The dataset is ideal for research that would benefit from an ensemble of equally-likely atmospheric states that are consistent with observations.

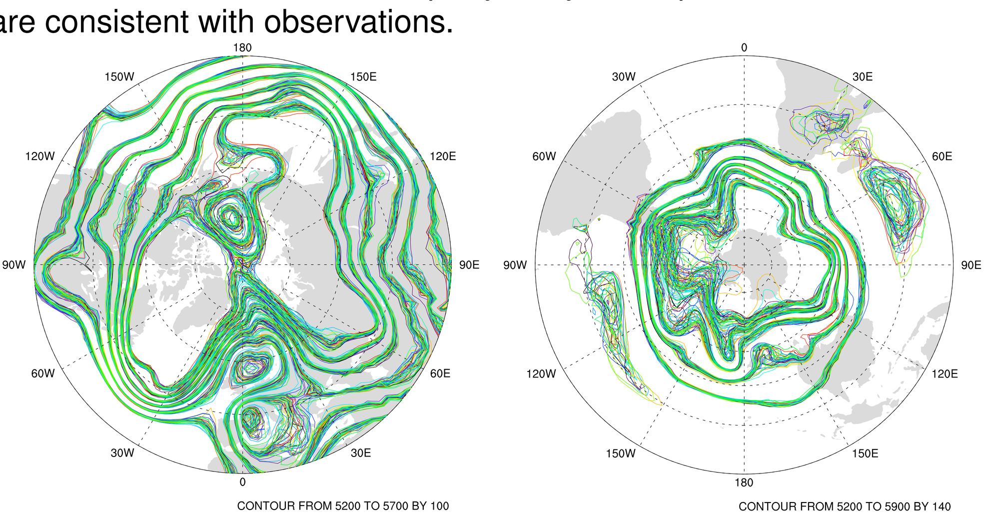


Figure 4: *Contours of the 500hPa geopotential height in 40 of the 80 CAM members for 6 hour forecasts valid 12:00 UTC 17 February 2003 (left) and 06:00 UTC 1 July 2001 (right). All of the model states are consistent with the observations, the ensemble captures the uncertainty.*

Assimilation Details

- 80 ensemble members of CAM Version 4
- "2-degree" (1.9° x 2.5°), 26 levels
- variables influenced by the assimilation: surface pressure, temperature, horizontal winds, specific humidity, cloud liquid, and cloud ice
- assimilation performed every 6 hours starting 1 Dec 1997
- globally, about 100,000 observations every 6 hours
- all members are forced by the same ocean analyses
- adaptive inflation used to maintain ensemble spread

CAM5 ("1-degree", 30 levels) has also been used in assimilations with DART, using the same setup as for CAM4. The results of both can be compared directly against the same observations.

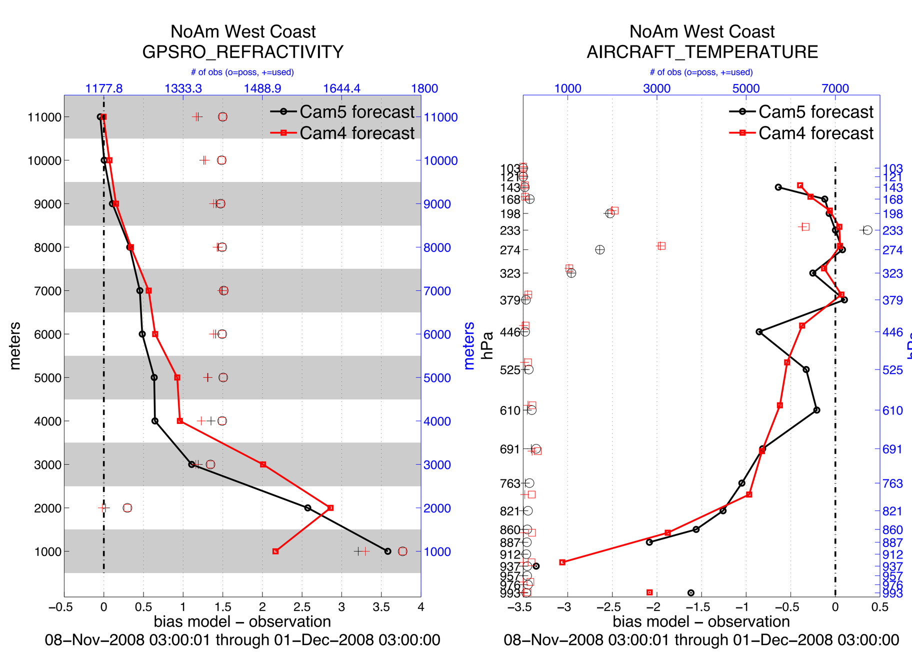


Figure 5: *The bias (bottom axis) of the default CAM4 (red) and CAM5 (black) models compared to Global Positioning System refractivity (left) and AIRCRAFT temperatures (right). The squares and 'plusses' show how many observations of that type were available and used, respectively (top axis). The domain covers longitudes 200-240 (160-120W) and latitudes 20N-60N: the North East Pacific.*

The time span of this comparison is only Nov. 7-30, 2008. Other seasons and years may have different bias characteristics. CAM5 compares favorably against CAM4 at most levels for the GPS observations. The large, apparent degradation at the lowest levels is puzzling, since CAM5 uses an improved boundary layer scheme. This layer may be too thick to represent the boundary layer exclusively. CAM5 has essentially the same bias as CAM4, relative to the AIRCRAFT observations. Most levels in that plot have a small number of observations against which to compare, which makes comparison of the curves less certain. More focused analyses of the 2 models can be conducted by using other observational data sets and refining the examined regions.

More details about CAM+DART can be found in Raeder, K., J. L. Anderson, N. Collins, T. J. Hoar, J. E. Kay, P. H. Lauritzen, and R. Pincus, DART/CAM: An Ensemble Data Assimilation for CESM Atmospheric Models J. Climate, 2012, DOI:10.1175/JCLI-D-11-00395.1 (in press).

4. Global Ocean Assimilation

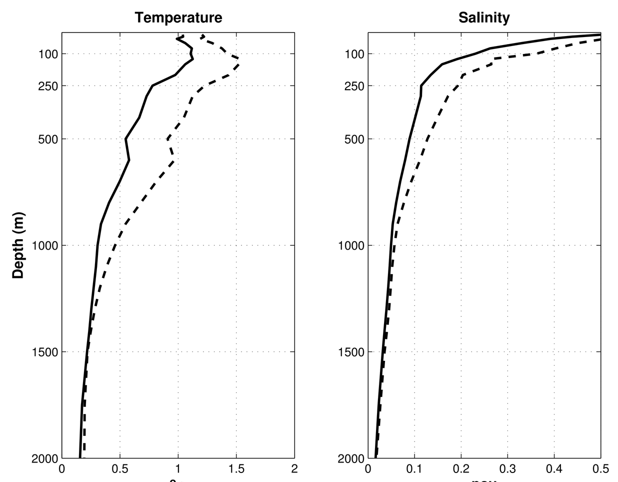
The CESM interfaces for the Parallel Ocean Program (POP) and the Community Land Model (CLM) support multiple instances, allowing data assimilation experiments to exploit unique atmospheric forcing for each POP or CLM model instance. This variety of atmospheric forcing is crucial for maintaining the ensemble spread of the ocean and land states, which improves the quality of the assimilations. To understand the role that data can play in constraining the ocean model, we compare the ensemble mean ocean analyses which include assimilation of WOD09 data ("Assim"), against the ensemble mean of ocean simulations without assimilation ("NoAssim").

POP Details

- 1-degree grid with displaced pole, 60 levels (POP gx1v6)
- 48 members initially drawn from a model climatology
- Atmospheric forcing for each POP member comes from a unique CAM ensemble member analysis (*e.g.* from Section 3)
- Prescribed sea ice concentration

Assimilation Details

- Use all temperature and salinity observations in the World Ocean Database 2009.
- Assimilate the observations in a +/- 12 hour window every midnight.



Assimilating subsurface temperature and salinity reduces the error of the model monthly average relative to the daily observations of temperature and salinity in the WOD09 data set.

Figure 6: *Average absolute error (misfit) between the daily observations of temperature (left) and salinity (right) from WOD09, and the model monthly average interpolated to the geographic location of the observations. Averages are taken over all observations at each depth level from 2000-2006. Solid line denotes the Assim model simulation and the dashed line is the the NoAssim simulation.*

The time-mean SST from 2000-2005 has more realistic western boundary currents in both basins, as shown by the similarity of the fields in Figure 7. The assimilation of data moves the SST field in a direction that reduces the mismatch between the wind-forced ocean SST ("NoAssim") and the observed satellite/in situ SST ("Hurrell").

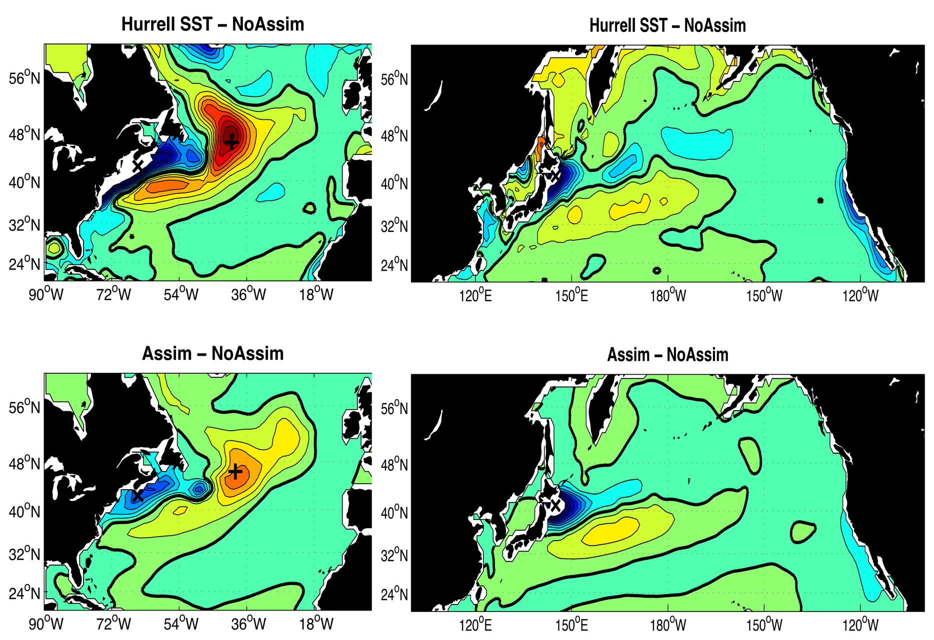


Figure 7: *Difference between the Hurrell SST (time average from 2000-2005) and the NoAssim SST (top) and difference between the Assim SST and the NoAssim SST (bottom) for the North Atlantic (left) and North Pacific (right). Contour intervals are $0.5^{\circ}C$, with the zero contour indicated in bold.*

5. Observation-Space Diagnostics

The performance of the assimilation is assessed by comparing the short-term forecast state to the observations about to be assimilated; a metric that is not dependent on a third party analysis. DART has a wide range of observation-space diagnostic tools to evaluate the performance of the assimilation.

DART has many more diagnostics than those shown here (Figures 5, 6 and 8); rank histograms, 3D plots of the observation locations color-coded to the observation value/QC value/rmse/bias/spread/rejected observations, mapping tools, . . .

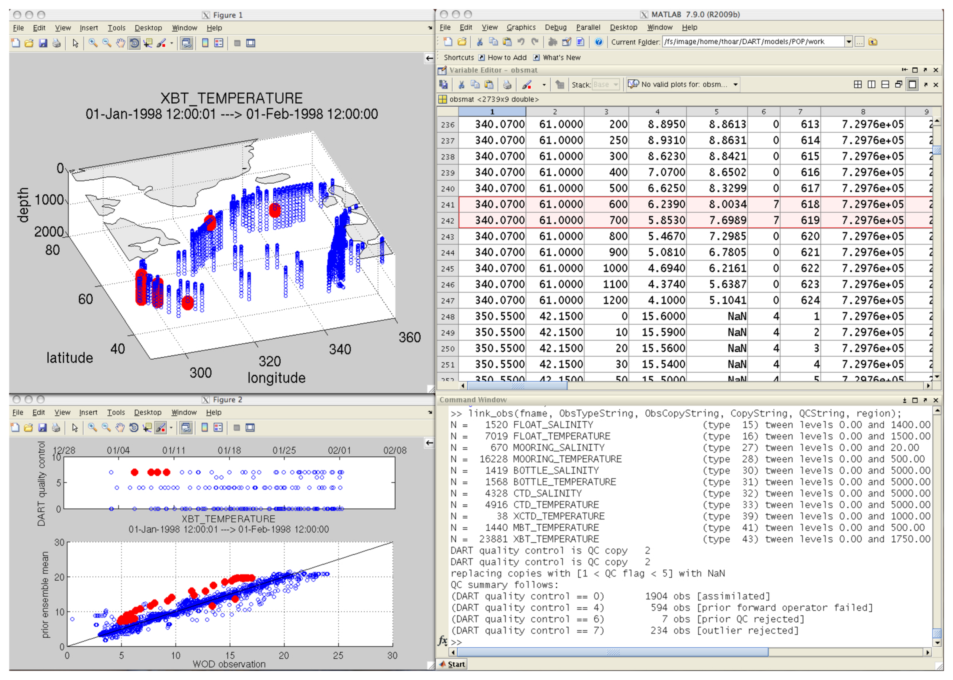


Figure 8: *DART's diagnostic tools make it easy to explore what observations are being rejected . . . and why. This is an example of some XBT observations in the North Atlantic. The information in the plots is linked — selecting observations in one view highlights them in all the views.*

6. Future CESM+DART Development

The DART interfaces will keep pace with developments in CESM, such as the adoption of the HOMME, cubed-sphere grid, and possibly the MPAS grid.

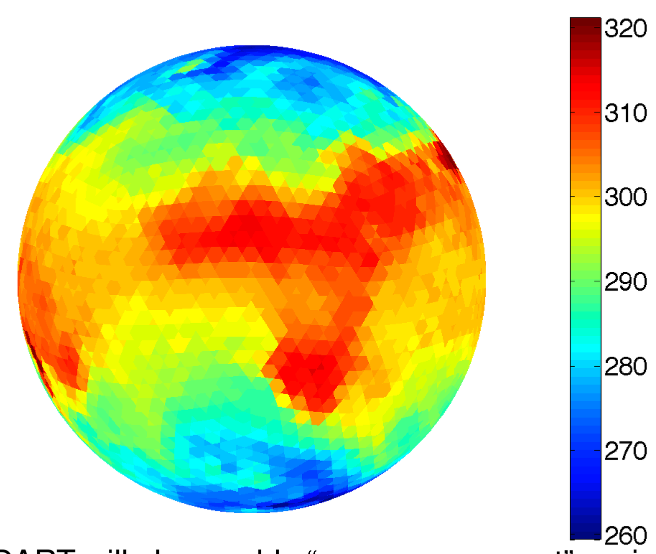


Figure 9: *Support for arbitrary grids is being extended. This is an MPAS temperature field from a DART diagnostic file.*

DART will also enable "cross-component" assimilation in CESM, in which an observation of the earth system can be assimilated into any component of CESM. For example, a sea-surface temperature observation might affect CAMs low level clouds.

7. Further Information

http://www.image.ucar.edu/DAReS/DART has information about downloading DART from our subversion server, a full DART tutorial (included with the distribution), and contacting us.

8. References

J. Anderson, T. Hoar, K. Raeder, H. Liu, N. Collins, R. Torn, and A. Arellano, 2009: *The Data Assimilation Research Testbed: A Community Data Assimilation Facility.* *BAMS* **90** No. 9 pp. 1283–1296

The National Center for Atmospheric Research is sponsored by the National Science Foundation.

The computational resources were provided by the Computational and Information Systems Laboratory at NCAR.