

Multi-instance CESM for Fully Coupled Data Assimilation Capability using DART



Kevin Raeder, Tim Hoar, Nancy Collins, Jeffrey Anderson: DAREs

Mariana Vertenstein, Tony Craig, Jim Edwards: SEWG

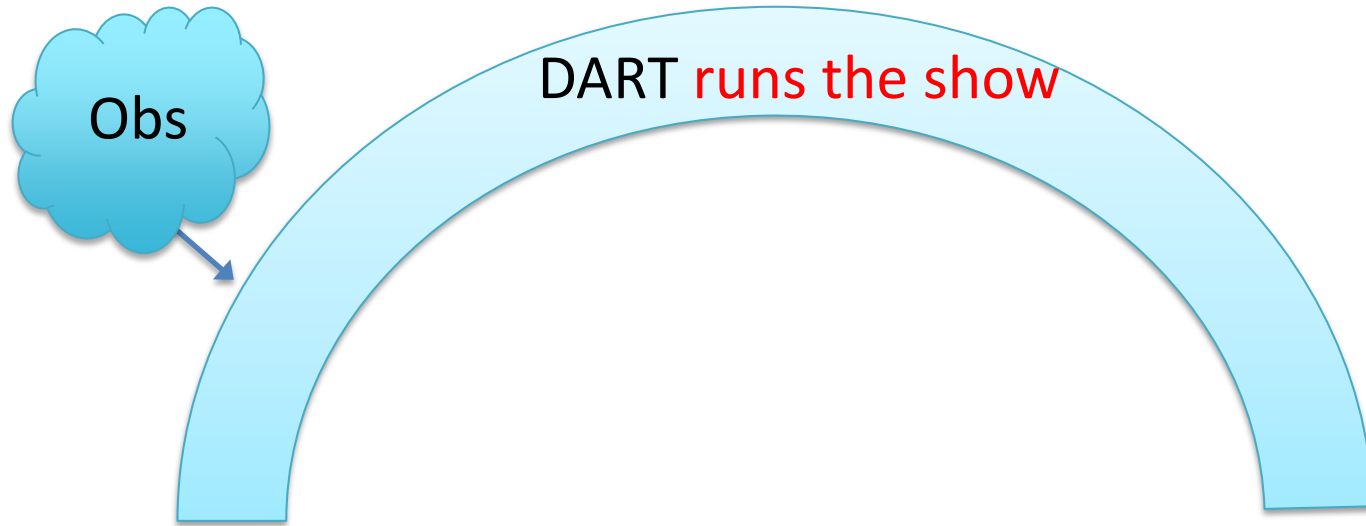
Alicia Karspeck: CGD/OS

Yongfei Zhang: U. of Texas Austin

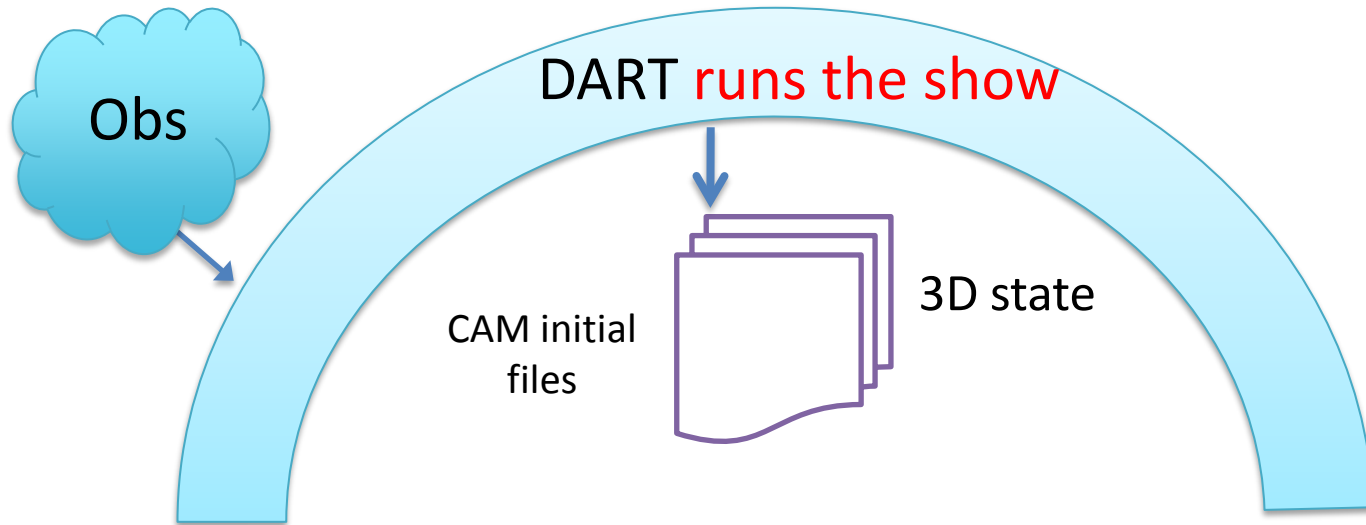
We are building an ensemble data assimilation system for NCAR's Community Earth System Model using DART (Data Assimilation Research Testbed)

- + Strong support from SEWG: Vertenstein, Craig, Edwards
- + Use of new multi-instance capability: CESM advances an ensemble of CAMs and/or POPs ... simultaneously.
- + CESM/CAM+DART is nearly as easy as CESM
- + Fully coupled data assimilation with any/all CESM components is within view.

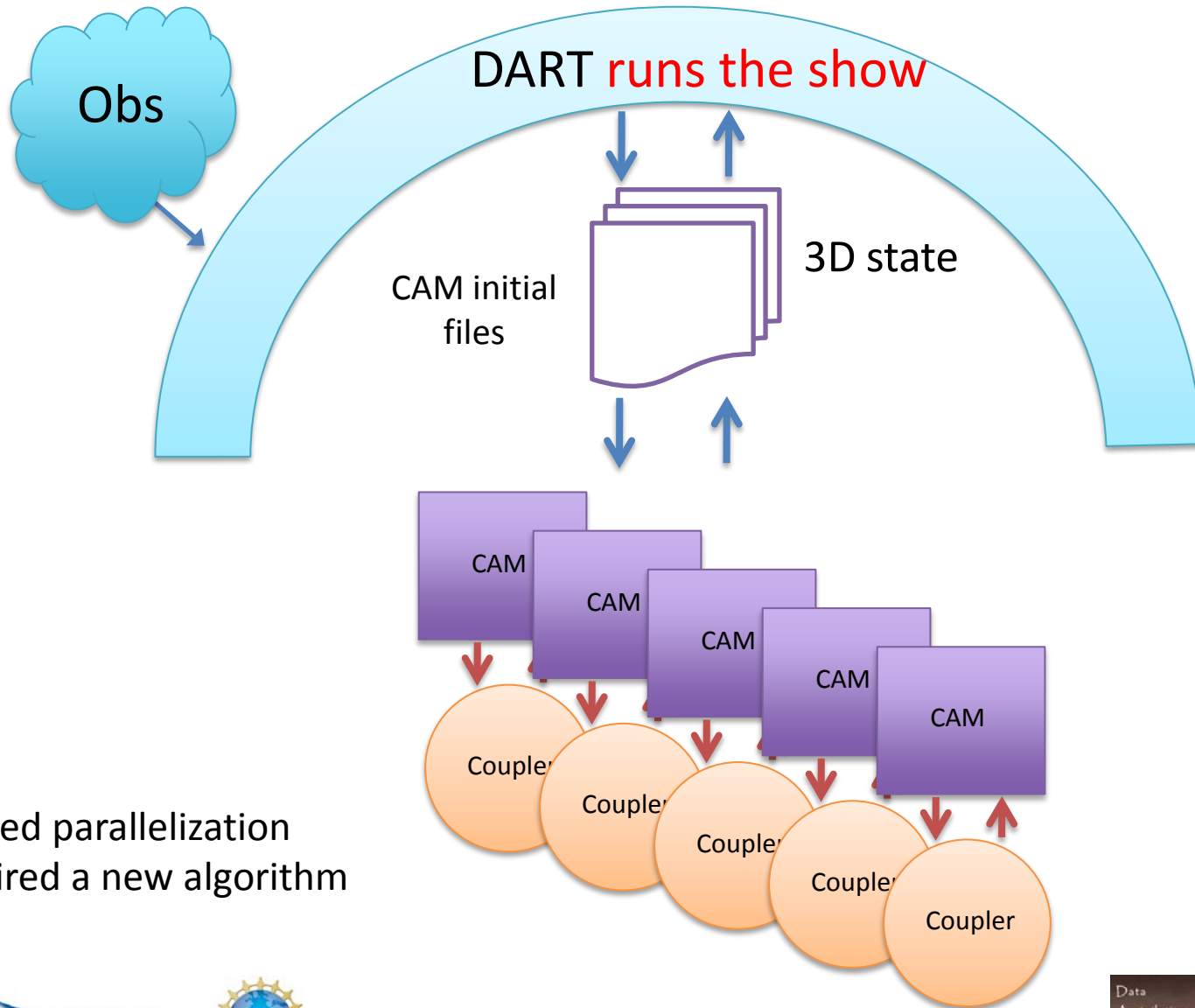
Past CAM Assimilation



Past CAM Assimilation



Past CAM Assimilation

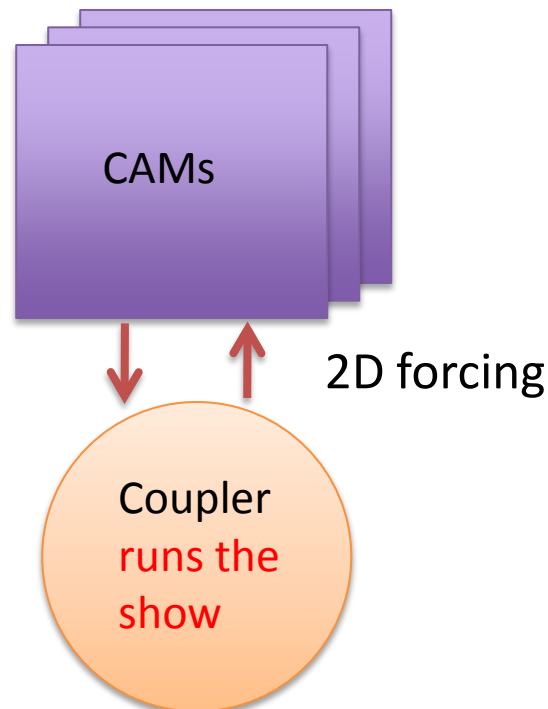


Limited parallelization
required a new algorithm

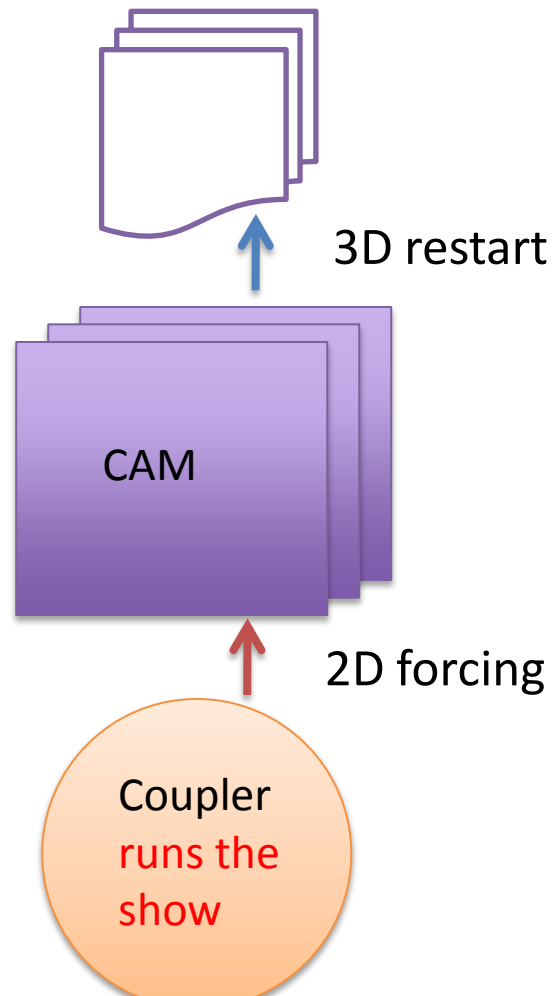


Current CAM Assimilation with CESM ensemble.

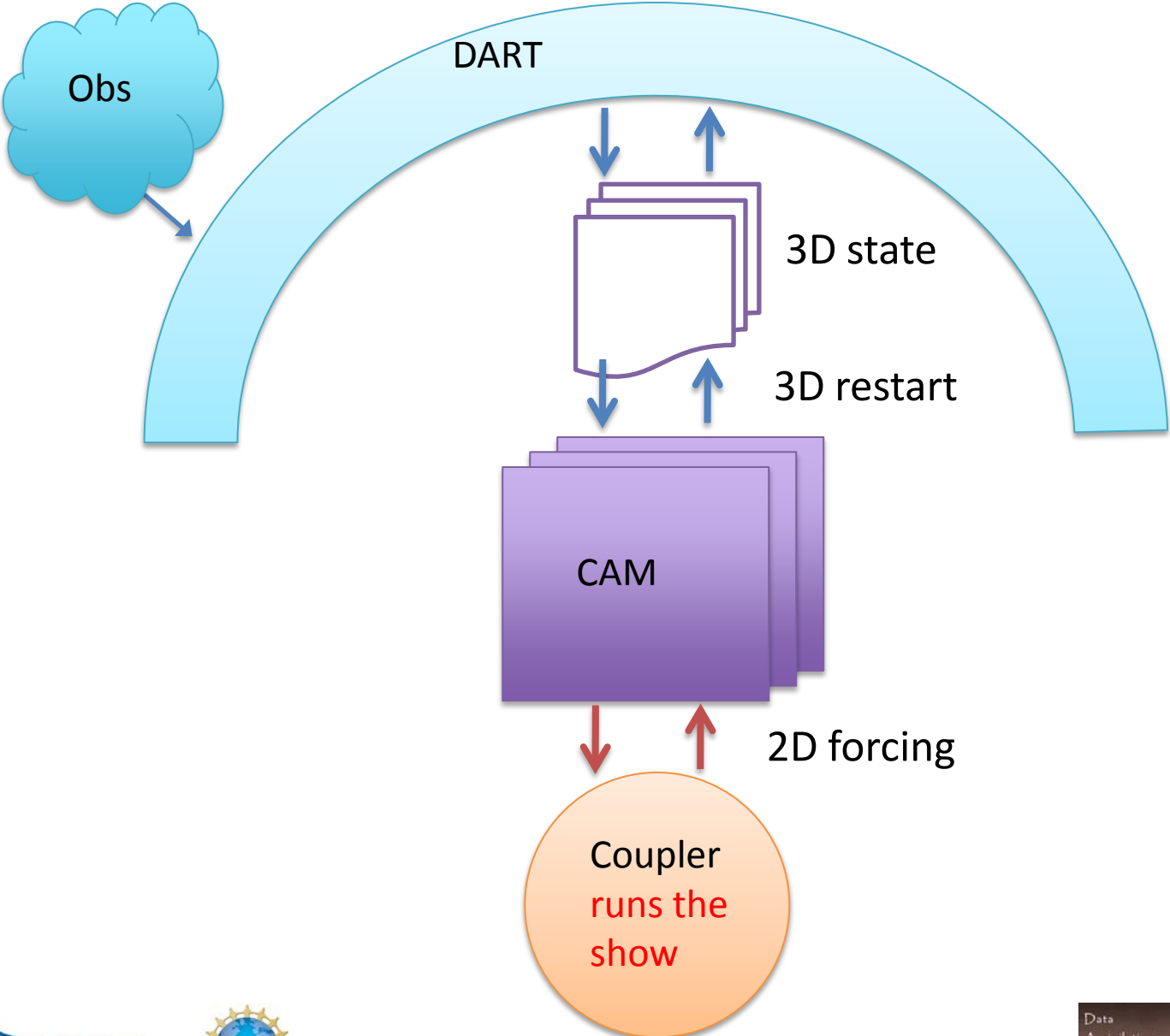
In a mind-boggling example of anticipating users' needs, SEWG was already working on the multi-instance capability and was looking for something to exercise it.



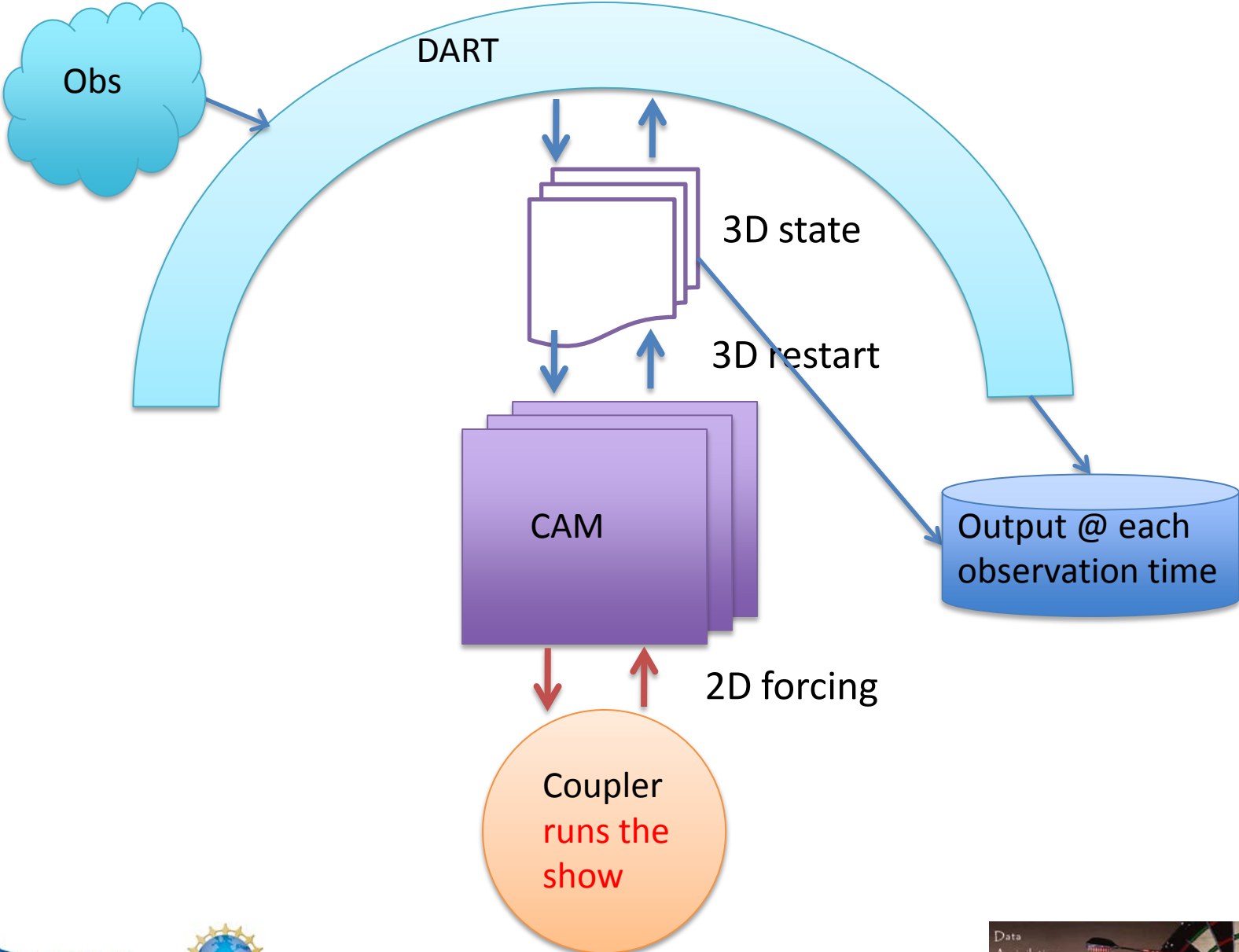
Current CAM Assimilation with CESM ensemble.



Current CAM Assimilation with CESM ensemble.



Current CAM Assimilation with CESM ensemble.



Advantages & Opportunities

- Any atmospheric model;
 - ▷ FV and Eulerian dy-cores, any resolution
 - ▷ CAM-MPAS interface; alpha testing nearing completion
 - ▷ CAM-SE interface is not done yet
 - ▷ Physics packages (CAM4, CAM5, ...)
 - ▷ WACCM; available, not tested, beta testers wanted
 - ▷ CAM+SKEBS (Berner), and other CAM variants
- Consistent with how the POP and CLM assimilations are being done
- Facilitates assimilation into a fully coupled model
- Advantages/applications we haven't thought of . . .

How hard is this really?

"I'm impressed by how easy you've made this to run " Robert Pincus, beta tester

How hard is this really?

"I'm impressed by how easy you've made this to run " Robert Pincus, beta tester

The CESM+DART setup script:

- ✓ Defines and builds a standard case (F_2000 for now),
- ✓ but uses the multi-instance capability for the atmospheric component.
- ✓ Modifies CESM scripts and namelists to
 - use namelist files appropriate for actual dates,
 - define the ensemble size,
 - manage the startup files,
 - run DART between the forecast and the archiving,
 - archive the new DART output.

How hard is this really?

"I'm impressed by how easy you've made this to run " Robert Pincus, beta tester

The CESM+DART setup script:

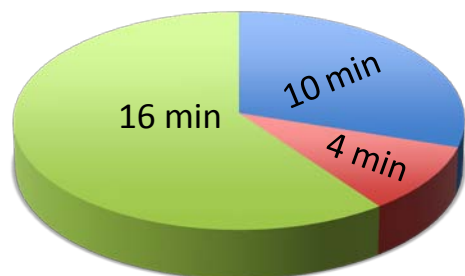
- ✓ Defines and builds a standard case (F2000 for now),
 - ✓ but uses the multi-instance capability for the atmospheric component.
 - ✓ Modifies CESM scripts and namelists to
 - use files appropriate for actual dates,
 - define the ensemble size,
 - manage the startup files,
 - run DART between the forecast and the archiving,
 - archive the new DART output,
- = A few hundred lines of code
- = *A few dozen lines which a user might want to change*

Validated on bluefire and hopper (NERSC Cray XT5).

Should run anywhere that CESM does.

Cost/Efficiency

1-deg CAM4+DART 6 hour Assimilation



- CESSM/CAM (x80)
- short-term archive
- DART

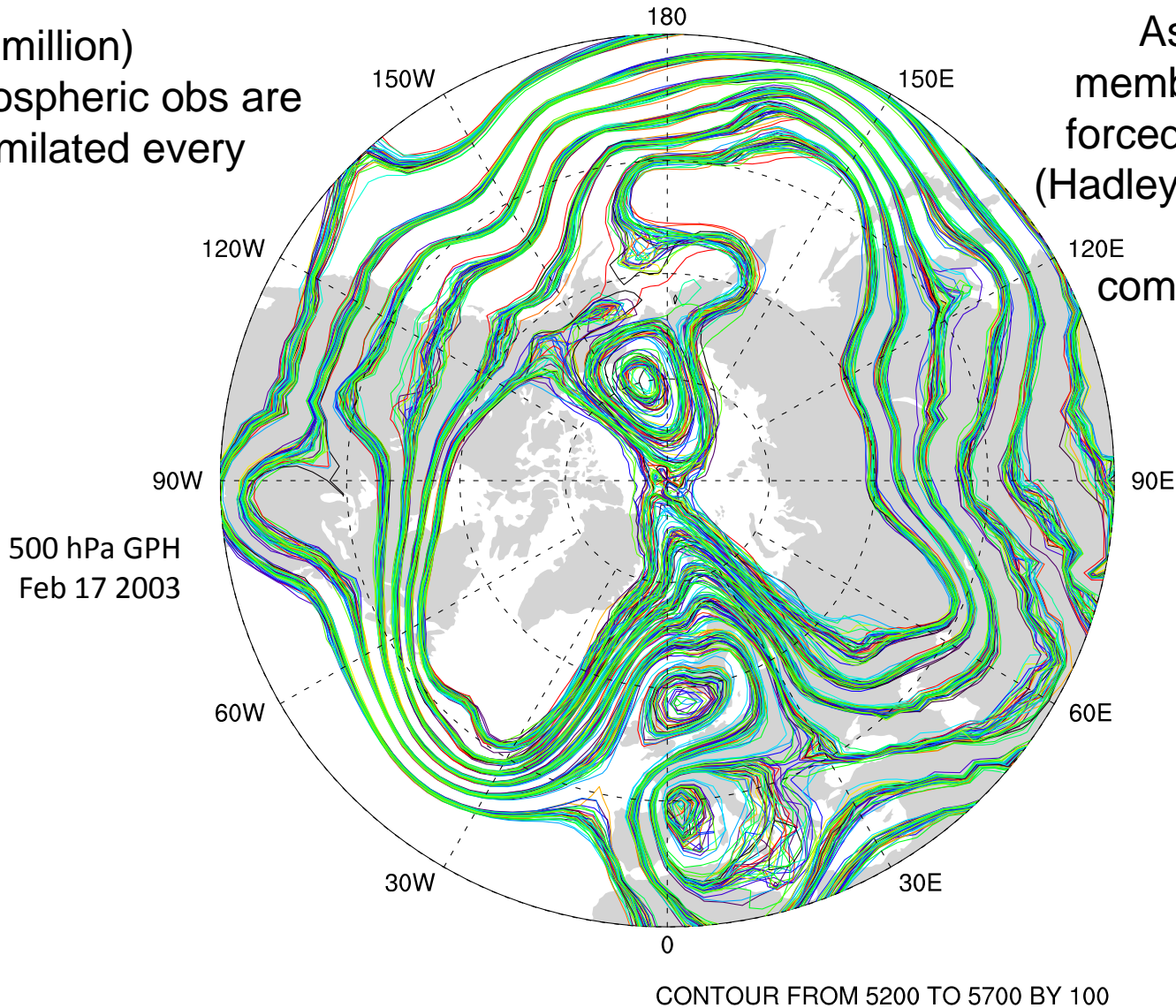
- 20 bluefire nodes × 2 hours/day
- in /ptmp (/glade is 10-100% slower)
- CAM5 has more levels, chemistry, larger state vector, ...
- Significant efficiencies are in the works
- and we get useful results from days, not decades,
- but yellowstone will be welcome! (we hope)

What do we get out of this?

Atmospheric Ensemble Reanalysis, 1998-2010

O(1 million)
atmospheric obs are
assimilated every
day.

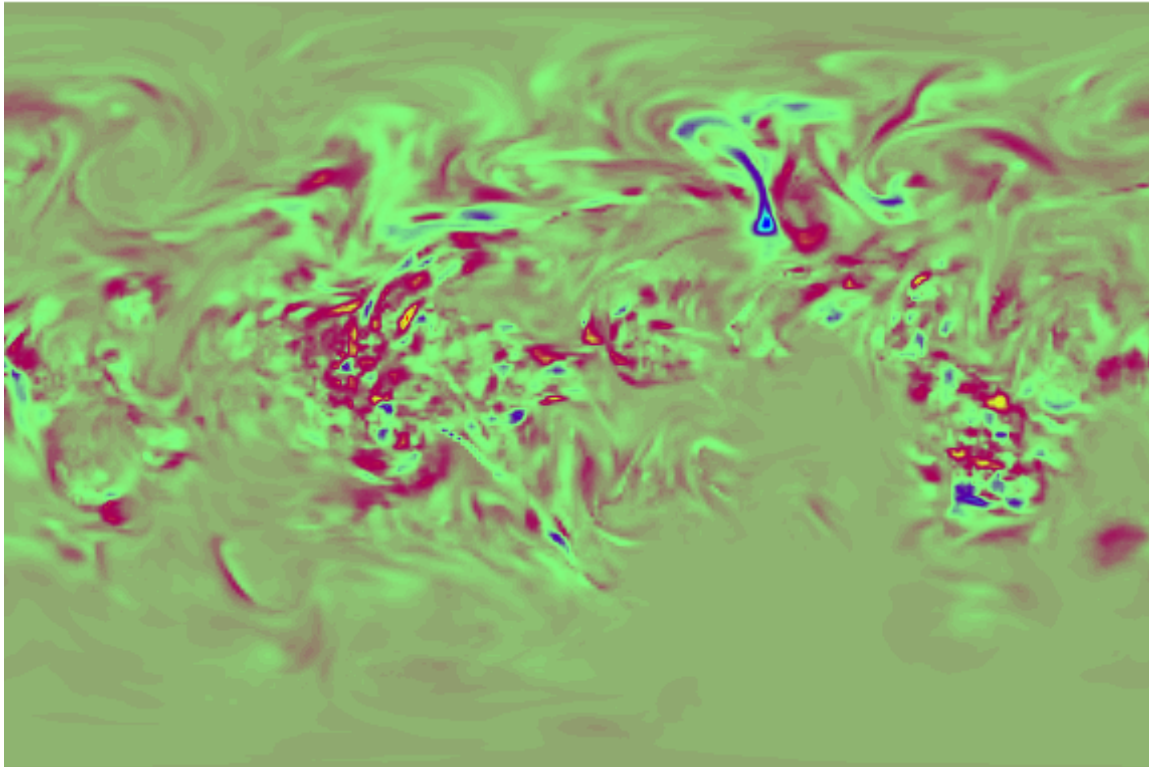
Assimilation uses 80
members of 2° FV CAM
forced by a single ocean
(Hadley+ NCEP-OI2) and
produces a very
competitive reanalysis.



Ensemble Mean Increment Due to Assimilation

These are some of the corrections to CAM's moisture resulting from assimilation of T, U, and V (no Q) observations.

Specific humidity

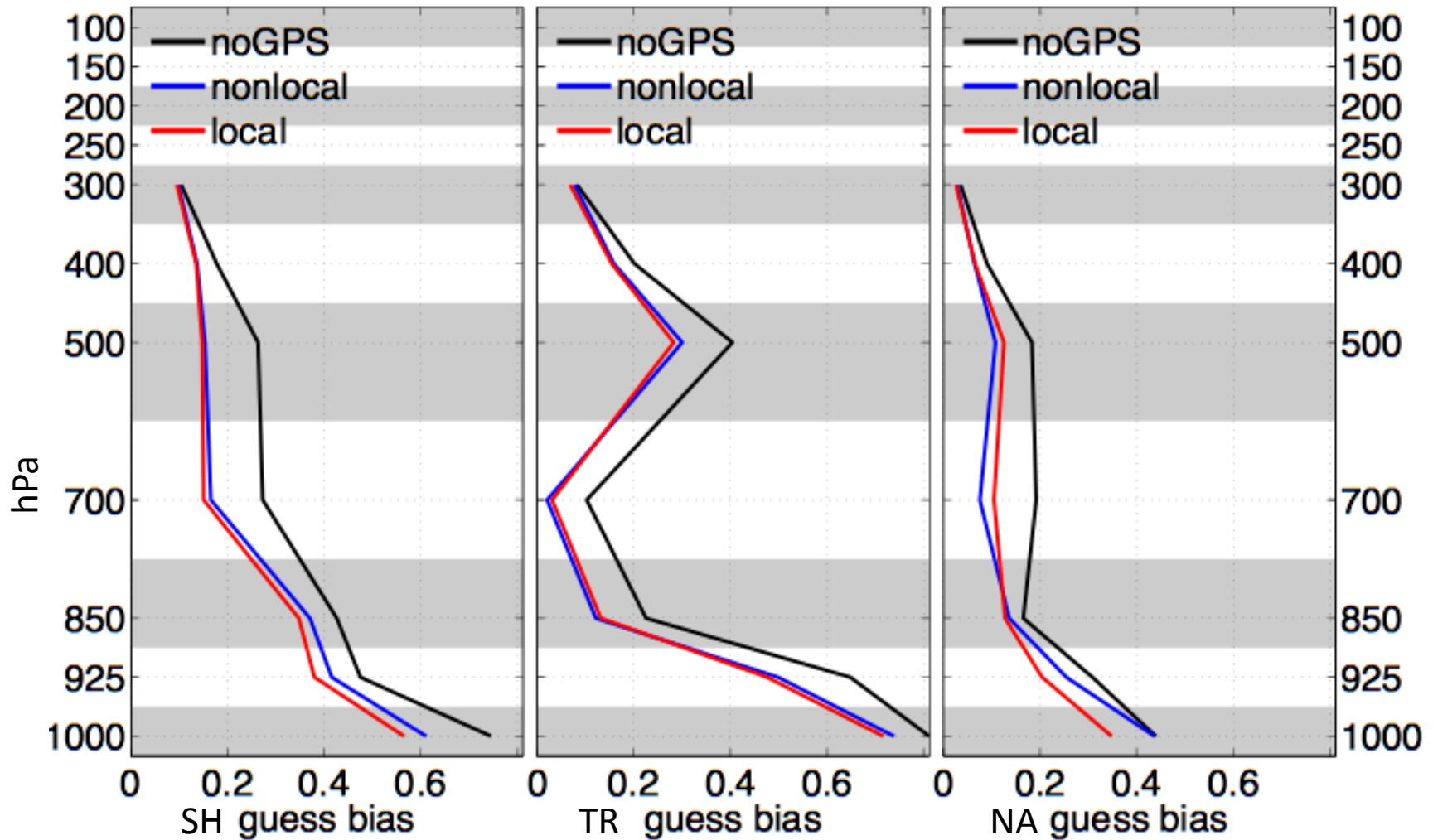


- ~ 200 hPa
- $-.06 < \Delta Q < .05$ g/kg
($-.05 < Q < .10$ g/kg)

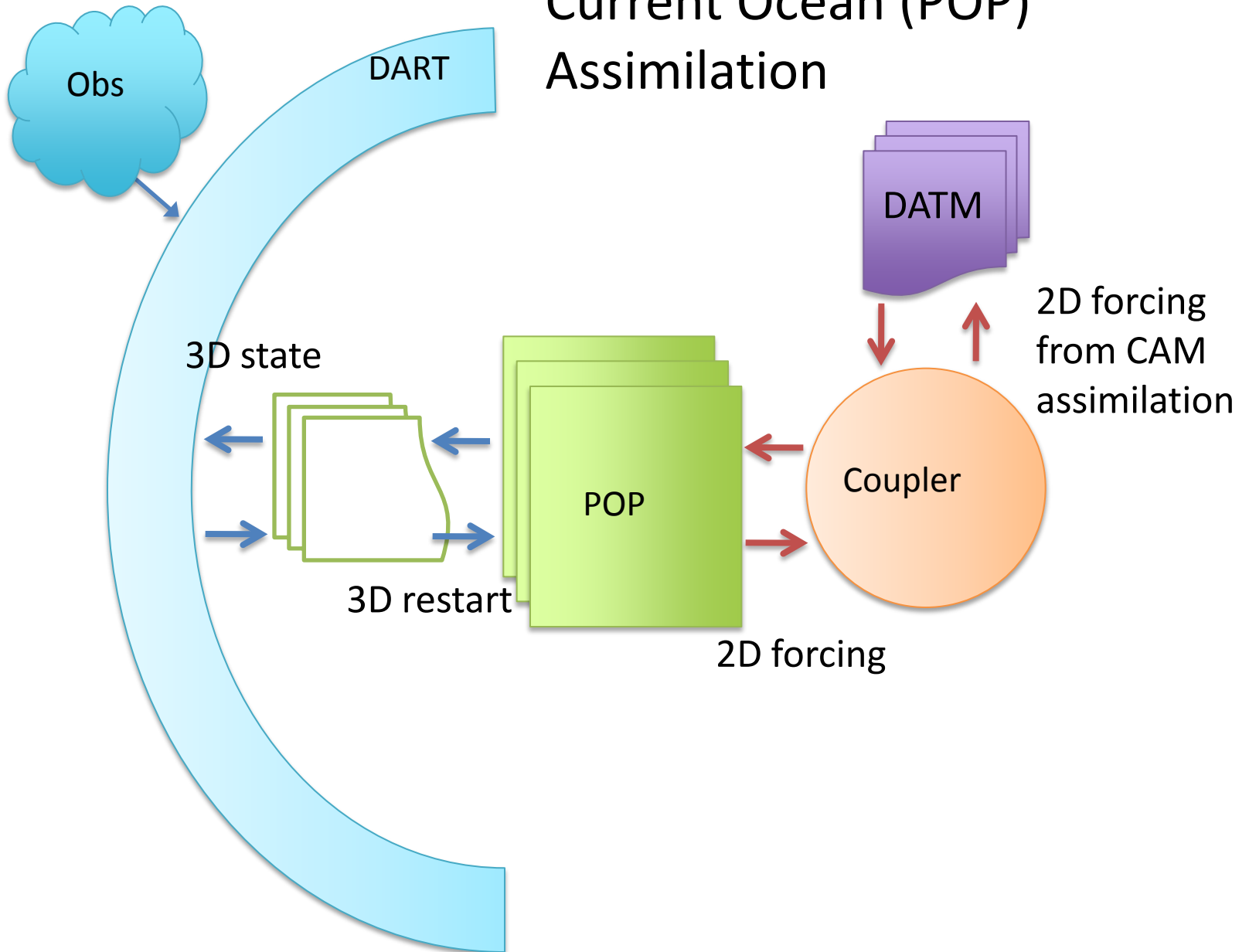


6-hour forecast Bias of Radiosonde Specific Humidity (Q)

December 2006

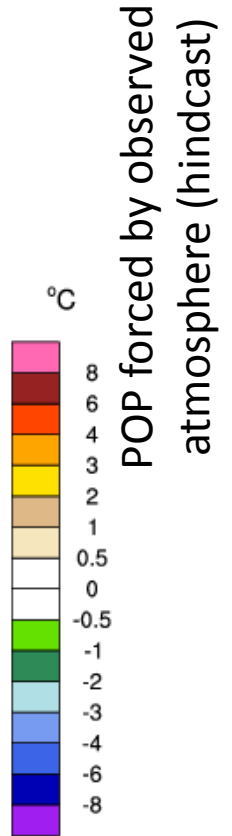
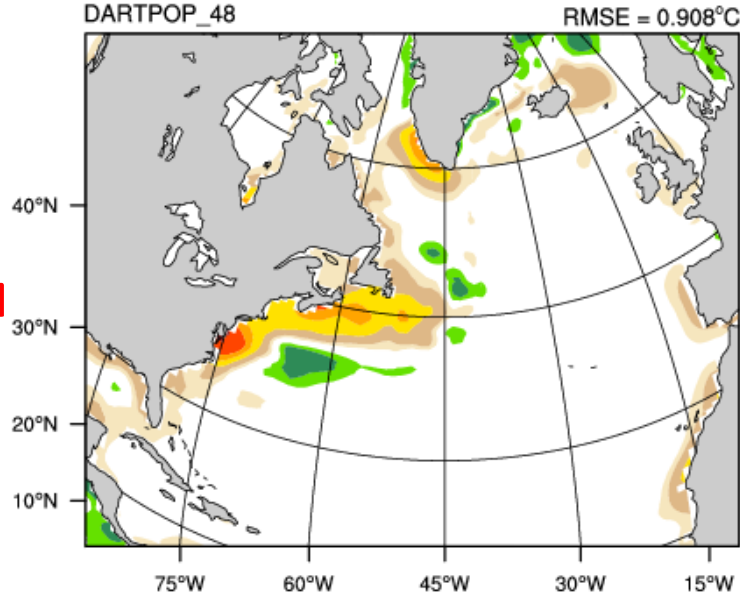
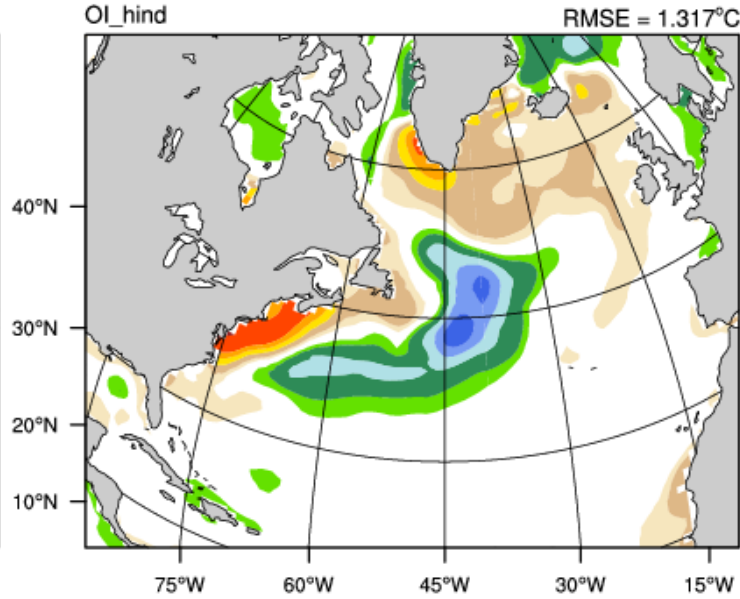
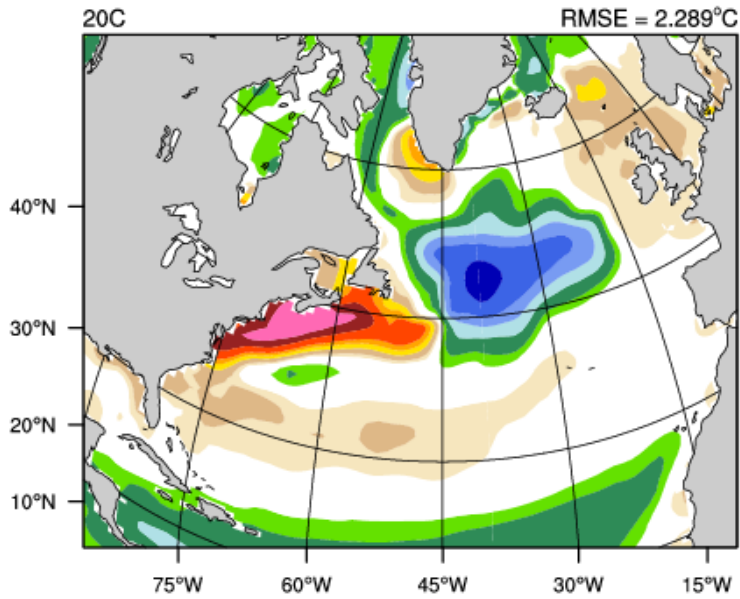


Current Ocean (POP) Assimilation



Physical Space: 1998/1999 SST Anomaly from HadOI-SST

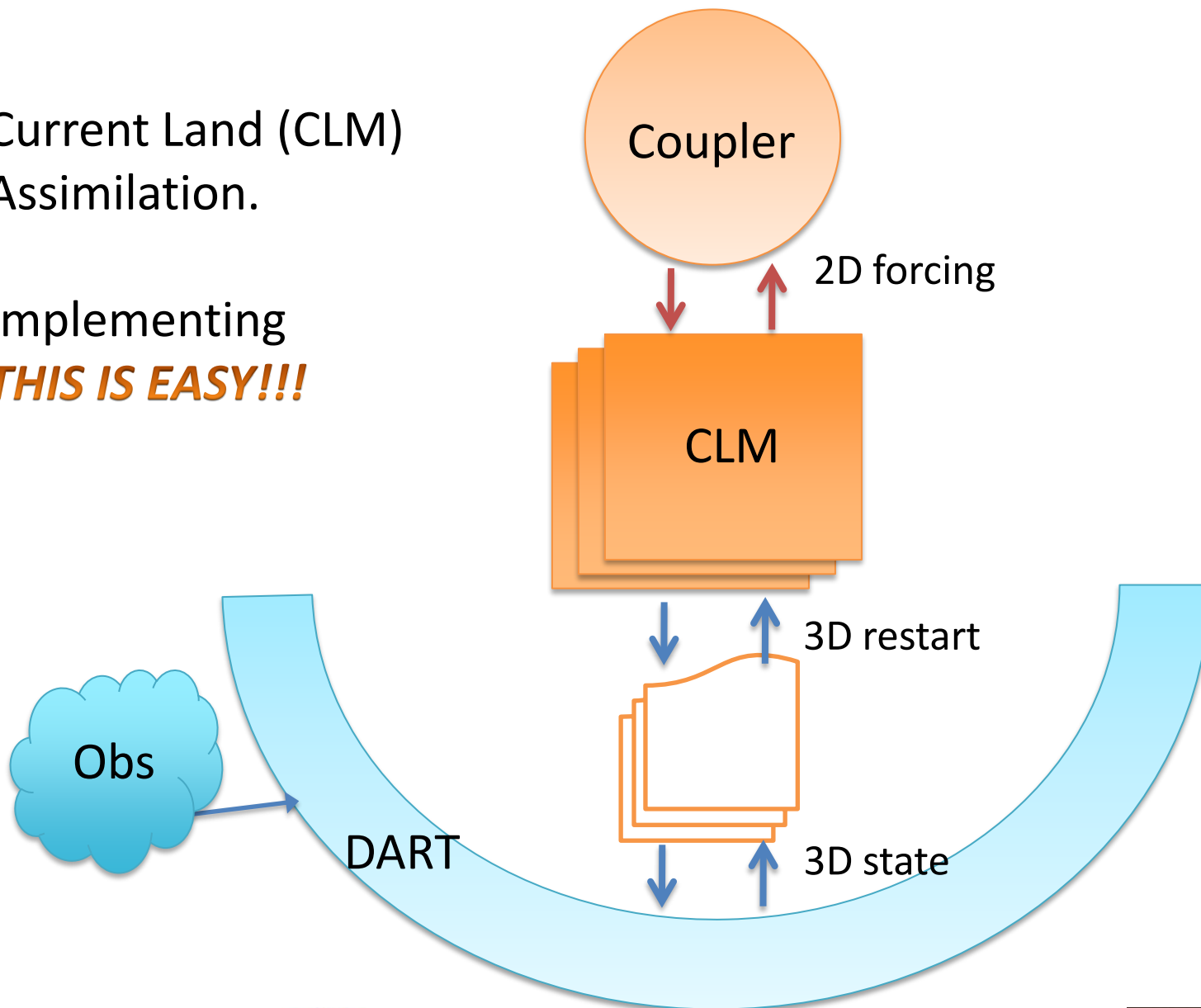
Coupled Free Run



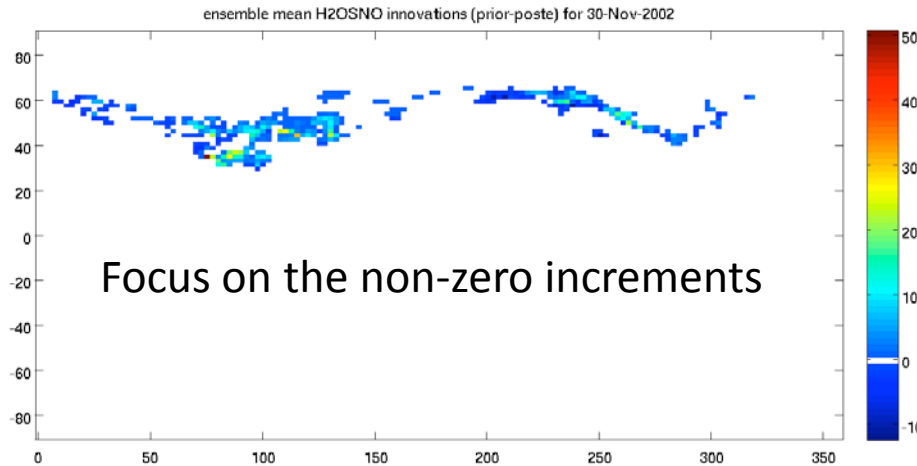
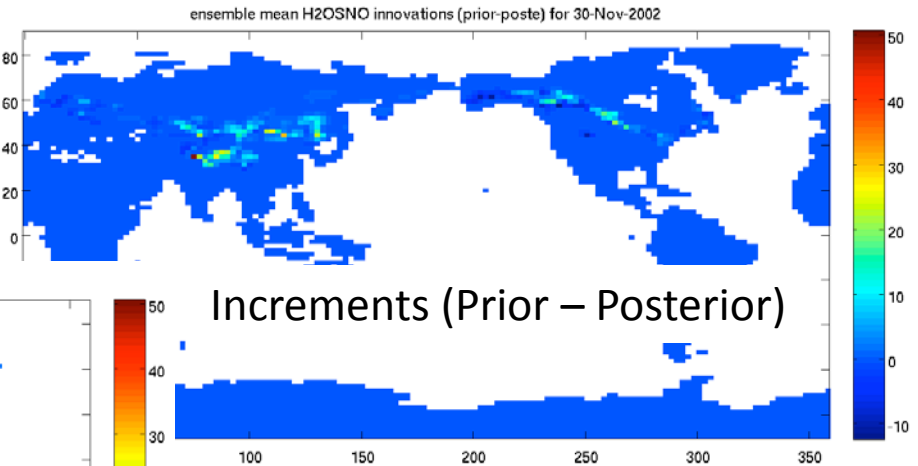
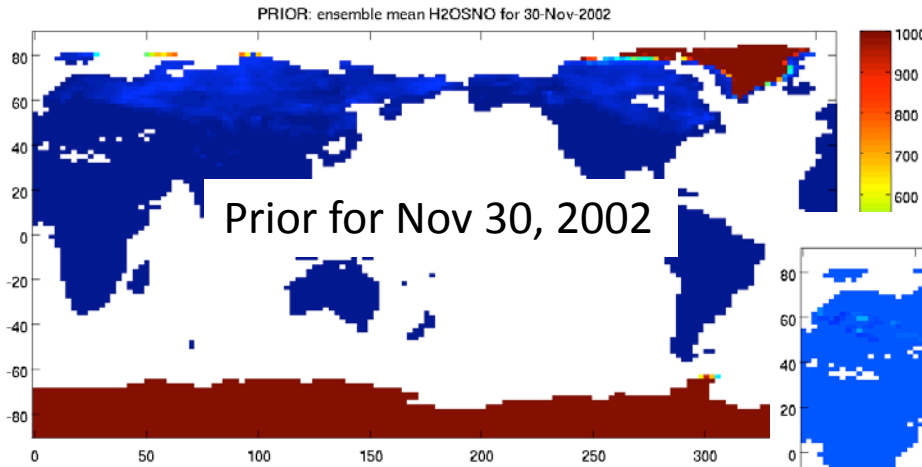
48 POP 48 CAM

Current Land (CLM)
Assimilation.

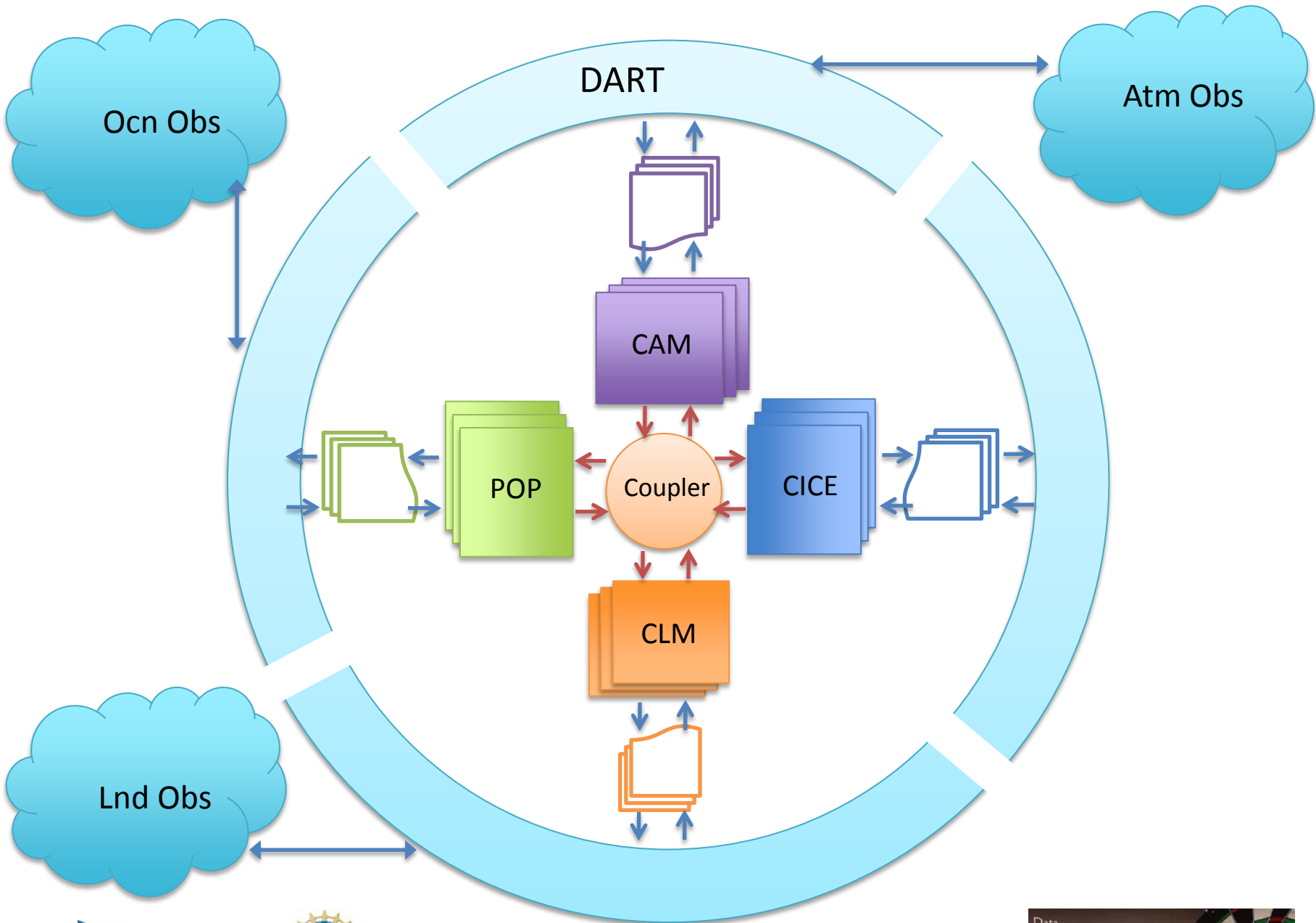
Implementing
THIS IS EASY!!!



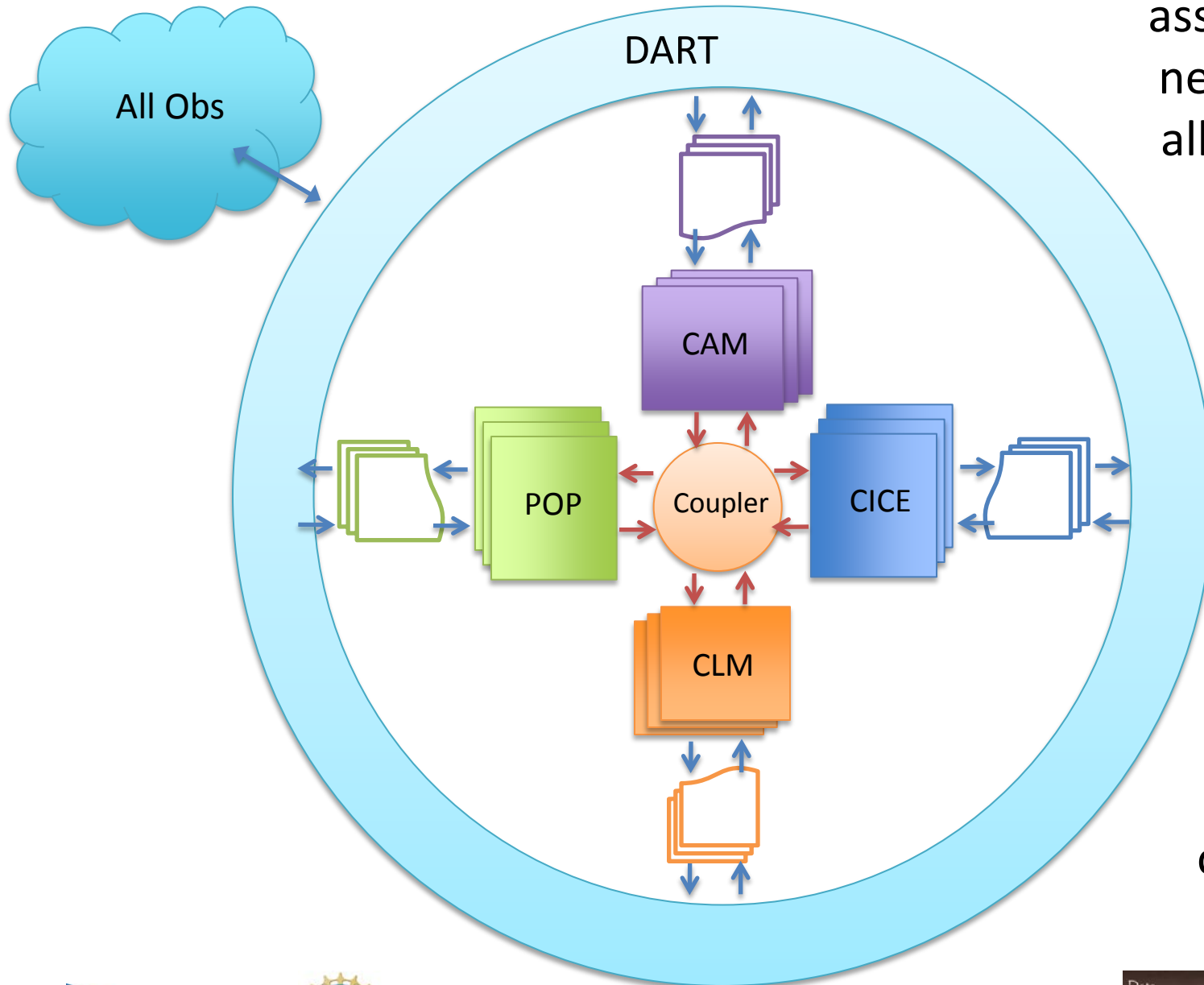
If we restrict ourselves to the simple cases ... here is the early result of an assimilation of MODIS snowcover fraction on total snow water equivalent in CLM.



The model state is changing in reasonable places, by reasonable amounts. At this point, that's all we're looking for.



Fully coupled assimilation will need data from all components at the same time



Each component corrected by all kinds of observations

Challenge for Earth System Model DA

- Atmospheric components of earth system models may not be as mature as NWP models.
- Model systematic or algorithmic errors may be large.
- Can lead to reduced quality analyses.
- But, DA can help to detect and correct errors!

Summary

CESM components+DART ensemble DA exists for:

- CAM: Multiyear ensemble reanalysis available,
- POP: Ensemble analyses used for decadal prediction initial conditions,
- CLM: Ensemble snow cover analyses and leaf area index research.

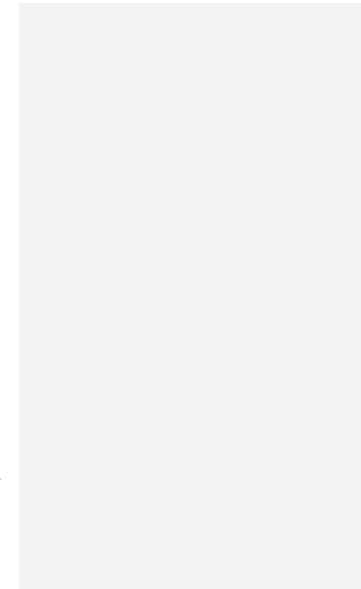
Plans and Challenges

- Build fully coupled earth system model ensemble assimilation system.
- Methods for dealing with land surface variables.
- Methods for dealing with strongly biased models.

Extras

Ensemble DA Sensitivity of Cyclone Central Pressure to Mean Sea Level Pressure (left) and 300 hPa geopotential height (right)

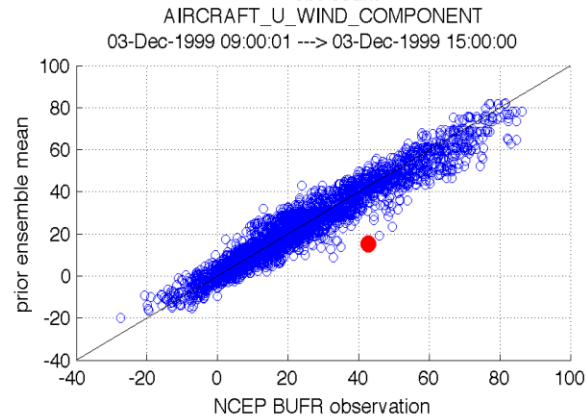
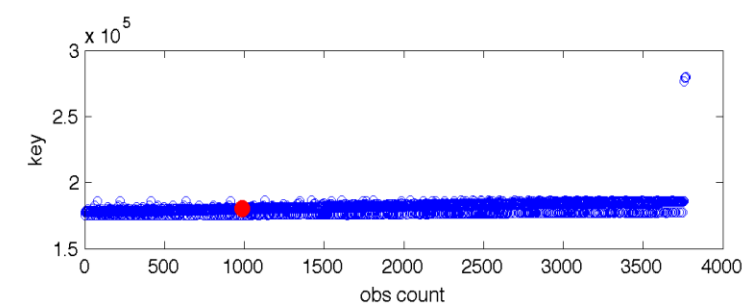
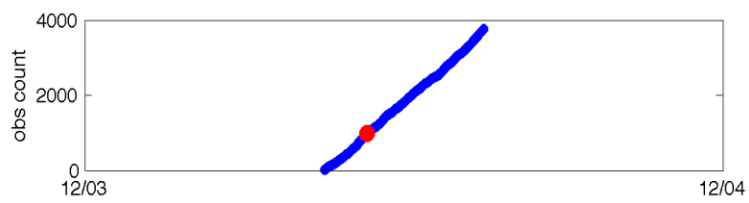
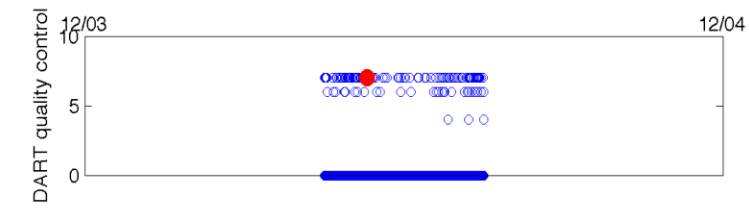
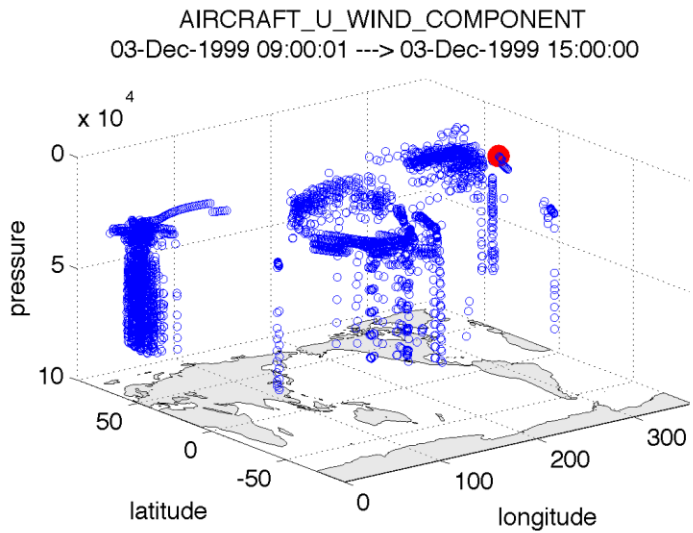
725
726 | Fig. 4-1: Sensitivity (in shades) of cyclone minimum pressure to MSLP (left panels) and 300 hPa
727 | z (right panels) for Case 1 (1981) 7.5-day forecasts. Contours represent ensemble mean forecast.



33

(Edmund Chang, et al., submitted to Monthly Weather Review 2011)

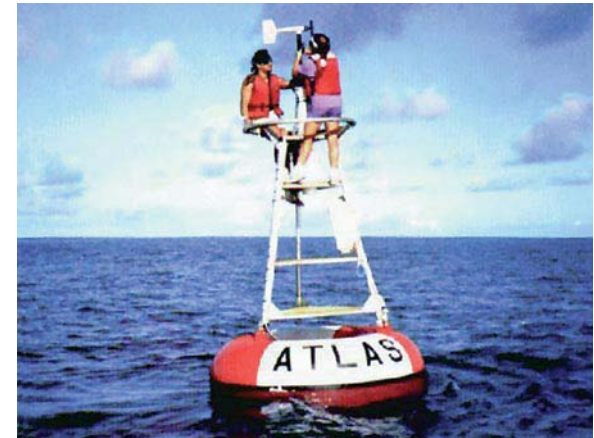
slide held in reserve



World Ocean Database 1,5 observation counts

These counts are for 1998 & 1999 and are representative.

FLOAT_SALINITY	68200
FLOAT_TEMPERATURE	395032
DRIFTER_TEMPERATURE	33963
MOORING_SALINITY	27476
MOORING_TEMPERATURE	623967
BOTTLE_SALINITY	79855
BOTTLE_TEMPERATURE	81488
CTD_SALINITY	328812
CTD_TEMPERATURE	368715
STD_SALINITY	674
STD_TEMPERATURE	677
XCTD_SALINITY	3328
XCTD_TEMPERATURE	5790
MBT_TEMPERATURE	58206
XBT_TEMPERATURE	1093330
APB_TEMPERATURE	580111



- temperature observation error standard deviation == 0.5 K.
- salinity observation error standard deviation == 0.5 msu.

The HARD part is: *What do we do when only SOME (or none!) of the ensembles have [snow,leaves,...] and the observations indicate otherwise?*

Corn Snow?

New Snow?

Sugar Snow?

Dry Snow?

Wet Snow?

“Champagne Powder”?

Slushy Snow?

Dirty Snow?

Early Season Snow?

Snow Density?

Sugar Snow?

Crusty Snow?

Old Snow?

Packed Snow?

Snow Albedo?



The ensemble **must** have some uncertainty, it cannot use the same value for all. The model expert must provide guidance. It's even worse for the hundreds of carbon-based quantities!

