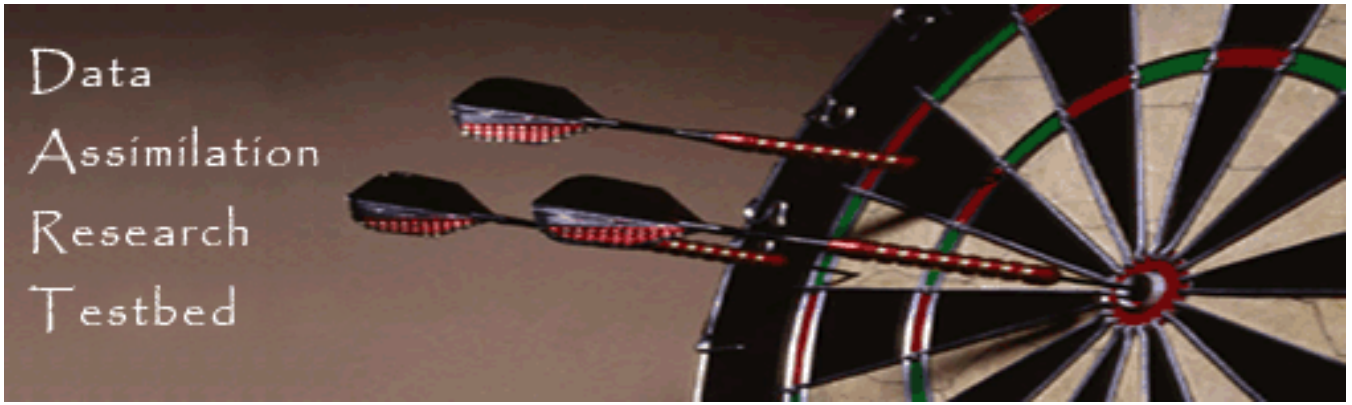
Multi-instance CESM for Fully Coupled Data Assimilation Capability using DART



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

Mariana Vertenstein, Tony Craig, Jim Edwards: SEWG

Jen Kay: CGD/AMPS

Alicia Karspeck: CGD/OS

Yongfei Zhang: U. of Texas Austin

Edmund Chang: Stony Brook U.

Why Assimilate Data?

- Observations are imperfect;
  - instrument error,
  - sparsity,
  - quantities we're not fundamentally interested in,
  - ...
- Models have errors in them, even if started from realistic ICs.
- Both have information

Extract the information from both and filter out the errors.

Ensemble Filter For Large Geophysical Models

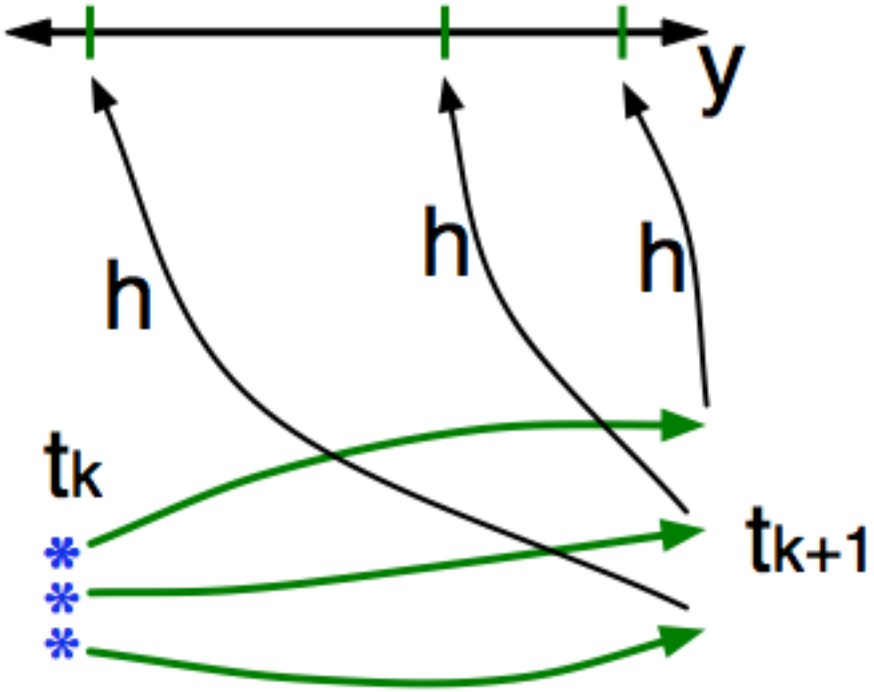
1. Use model to advance ensemble (3 members here) to time at which next observation becomes available.

Ensemble state estimate, $x(t_k)$, after using previous observation (analysis)

Ensemble state at time of next observation (prior)

Ensemble Filter For Large Geophysical Models

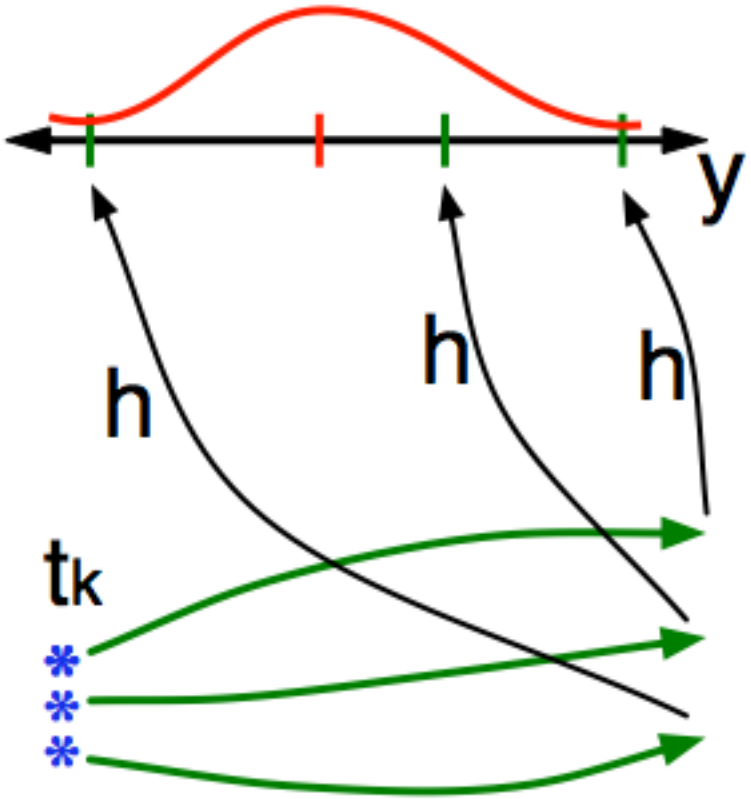
2. Get prior ensemble sample of observation, y = h(x), by applying forward operator **h** to each ensemble member.



Theory: observations from instruments with uncorrelated errors can be done sequentially.

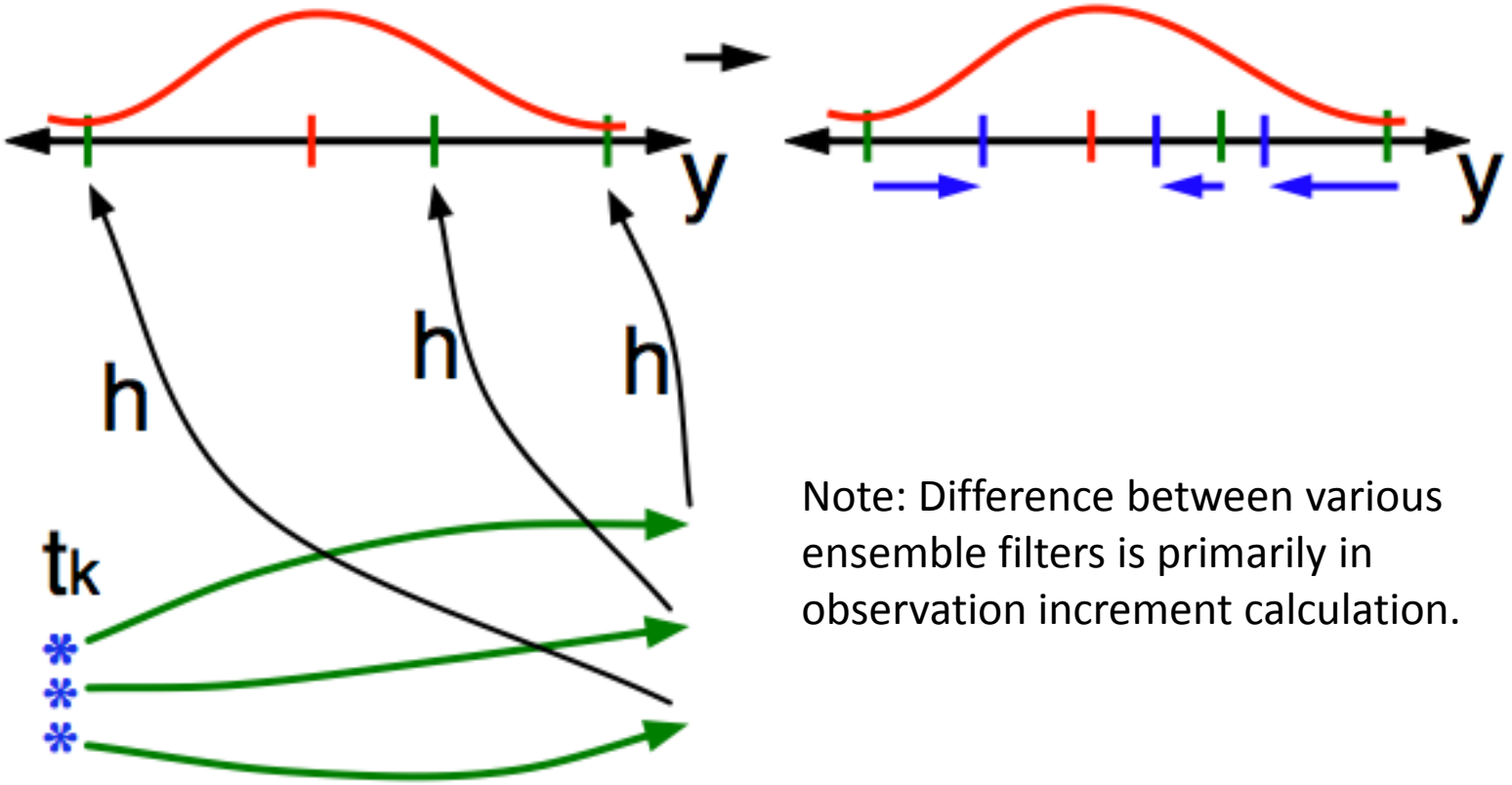
Ensemble Filter For Large Geophysical Models

3. Get observed value and observational error distribution from observing system.



Ensemble Filter For Large Geophysical Models

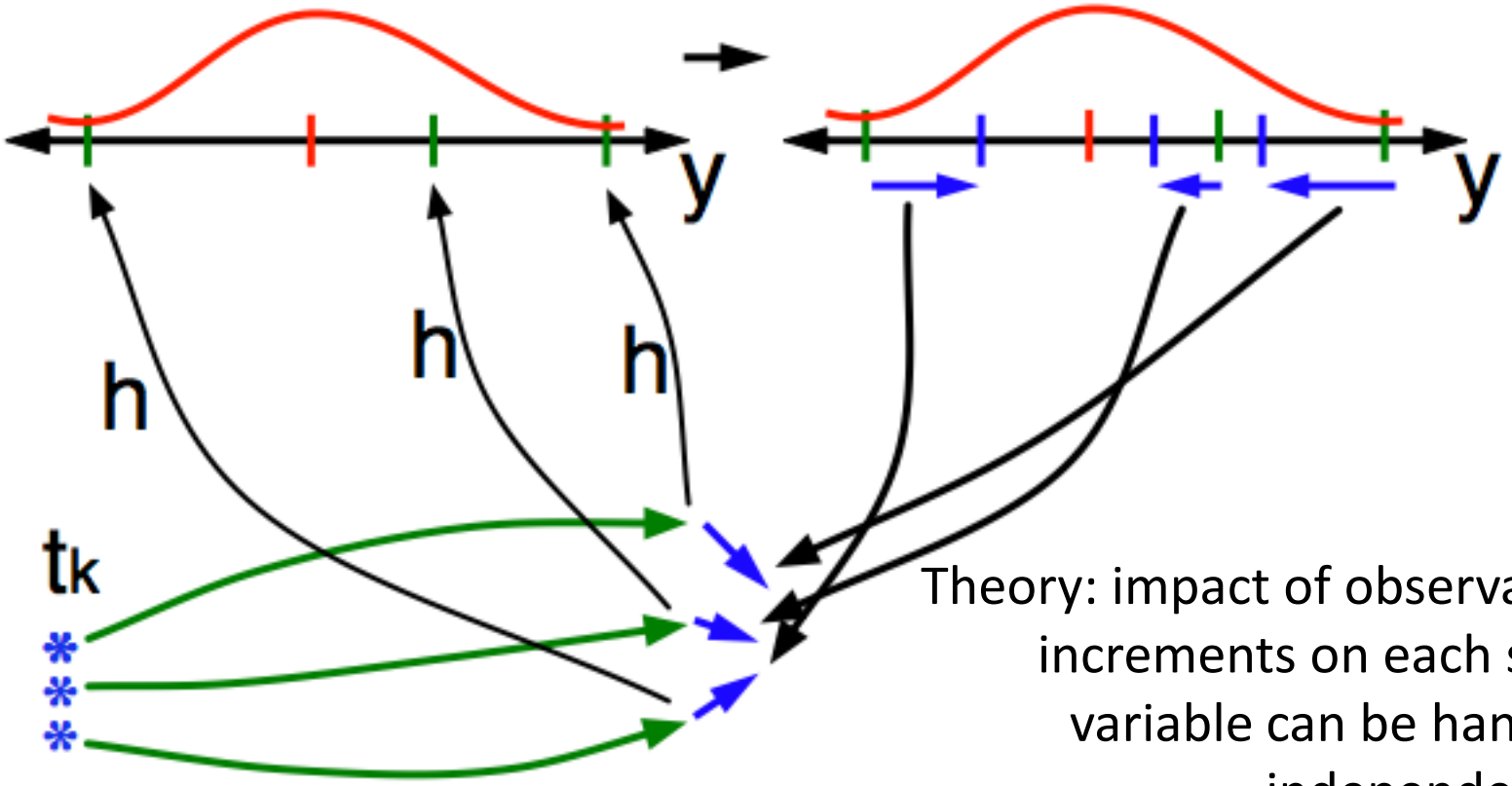
4. Compute the increments for the prior observation ensemble (this is a scalar problem for uncorrelated observation errors).



Note: Difference between various ensemble filters is primarily in observation increment calculation.

Ensemble Filter For Large Geophysical Models

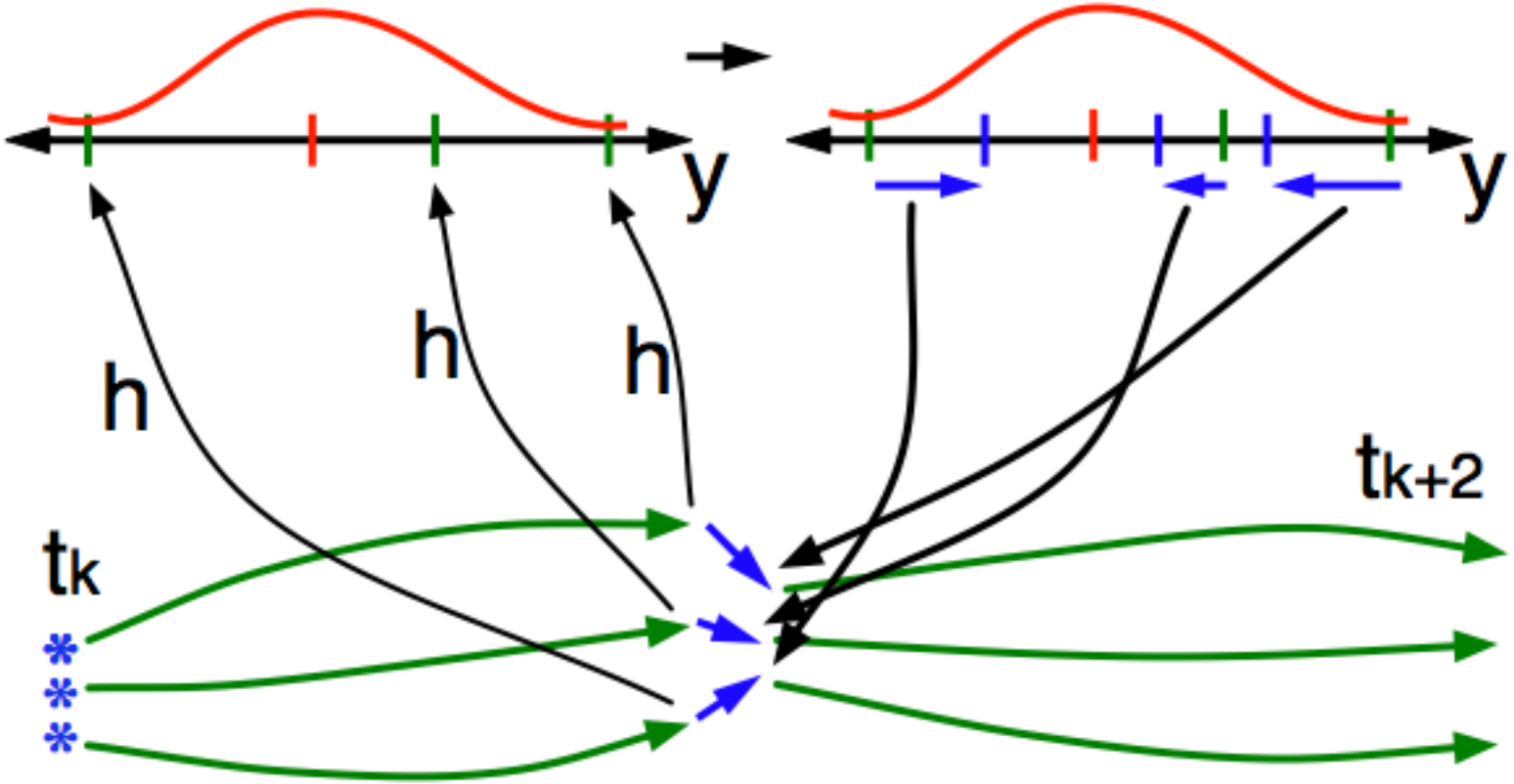
5. Use ensemble samples of **y** and each state variable to linearly regress observation increments onto state variable increments.



Theory: impact of observation increments on each state variable can be handled independently!

Ensemble Filter For Large Geophysical Models

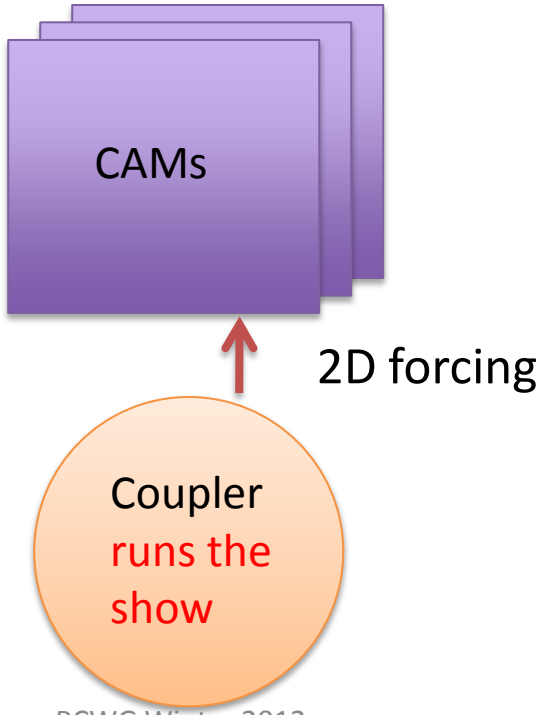
6. When all ensemble members for each state variable are updated, there is a new analysis. Integrate to time of next observation …



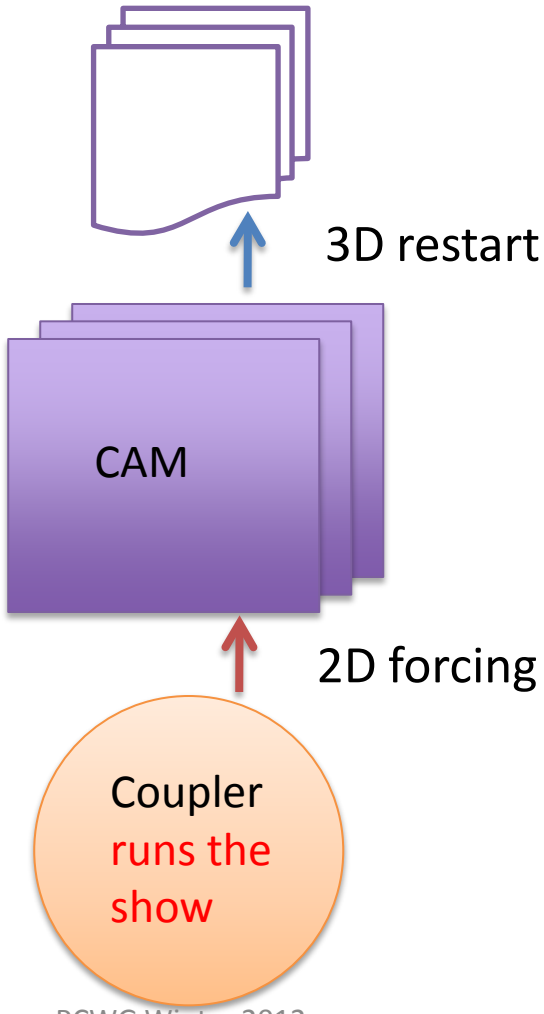
We are building an ensemble data assimilation system for NCAR's Community Earth System Model using DART (Data Assimilation Research Testbed)

+ 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.

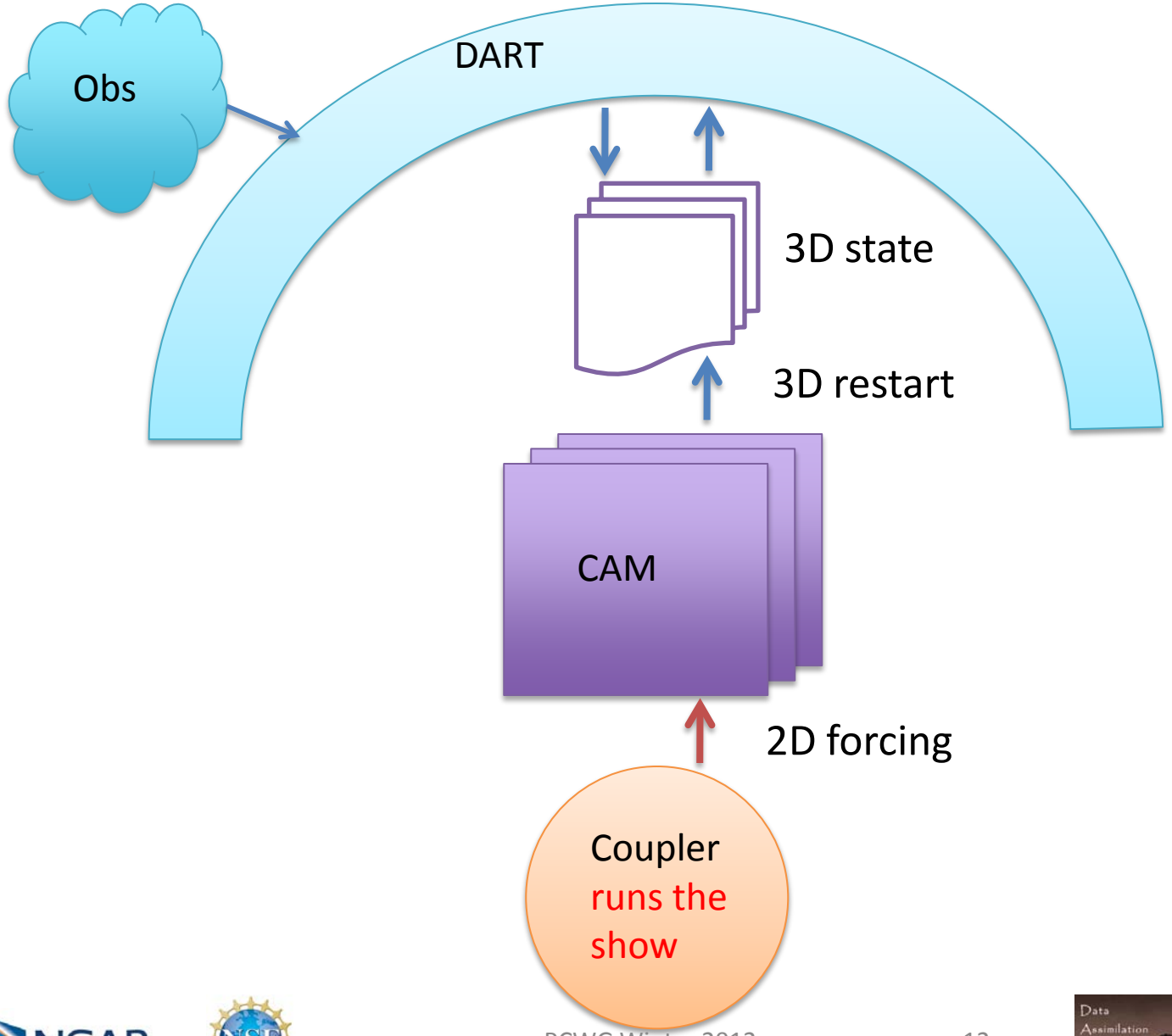
Current CAM Assimilation with CESM ensemble.



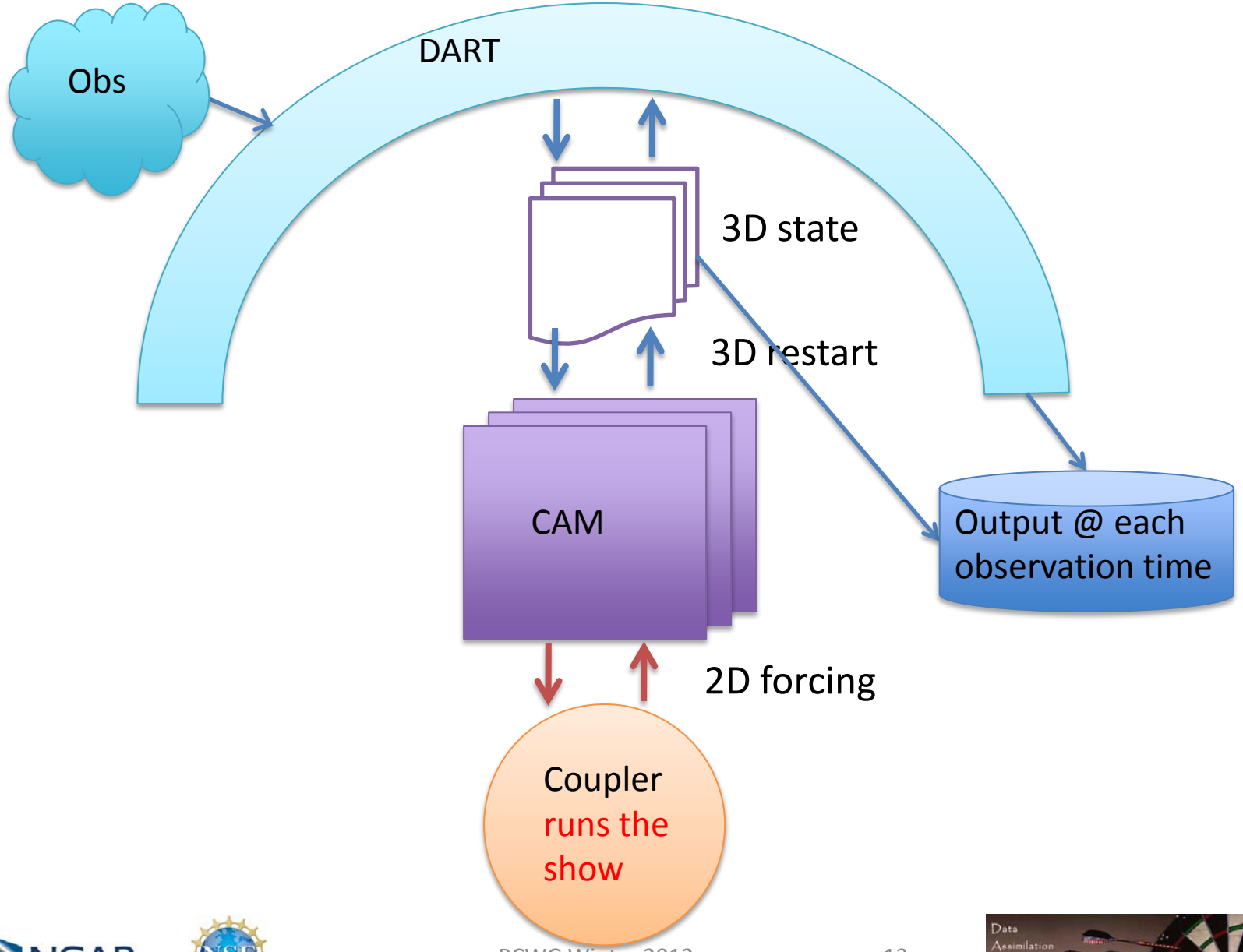
Current CAM Assimilation with CESM ensemble.



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,

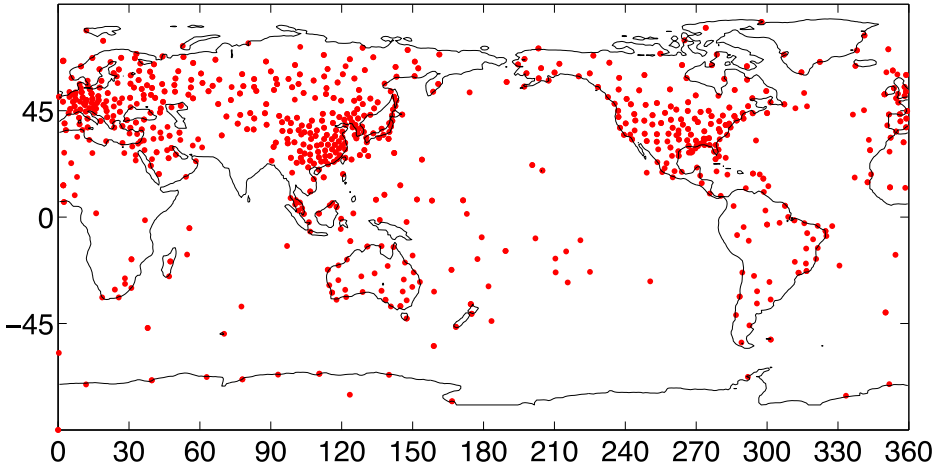
= A few hundred lines of code

= A few *dozen* lines which a user *might* want to change

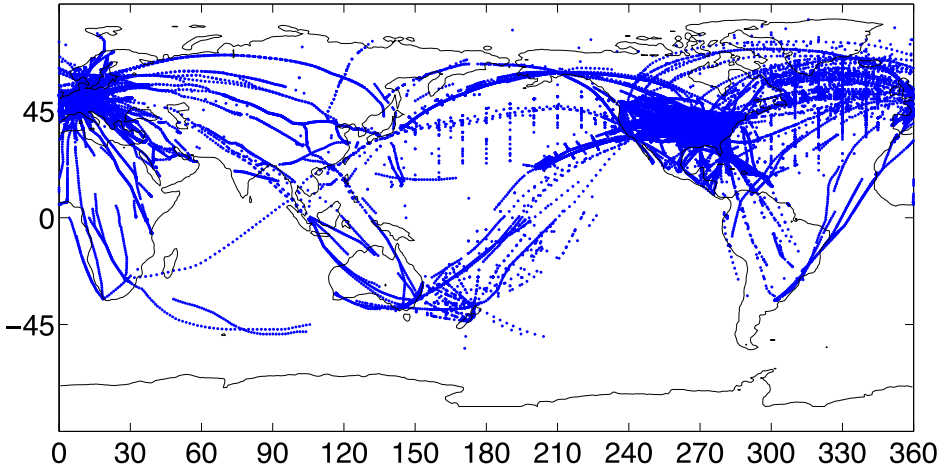
Validated on bluefire and hopper (NERSC Cray XT5).
Should run anywhere that CESM does.

Typical daily atmospheric observation set coverages (e.g. 12/6/2006)

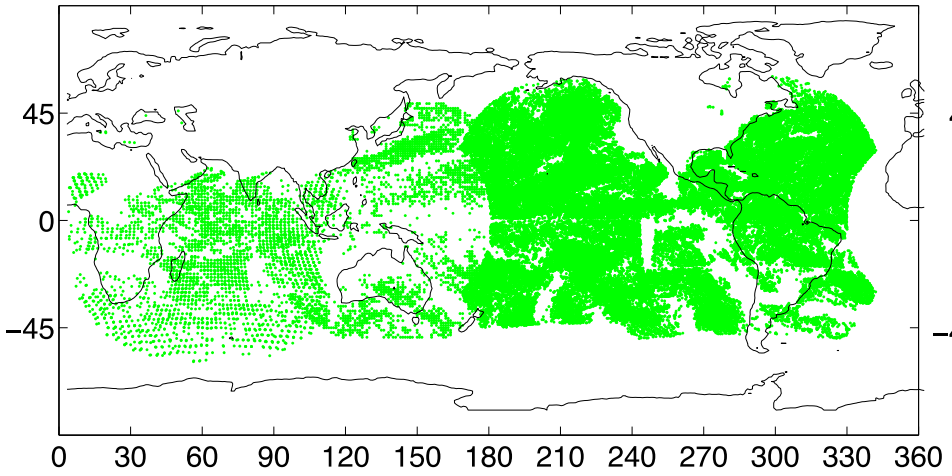
- Radiosonde



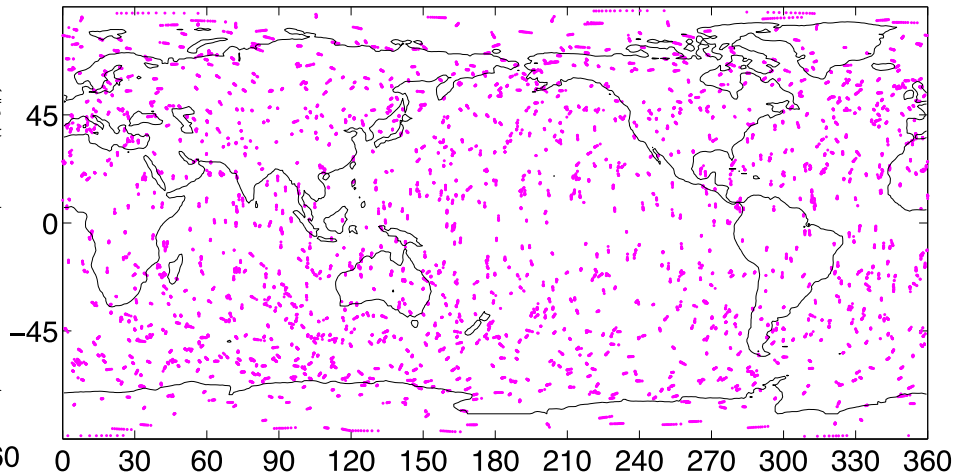
- ACARS/Aircraft



- Satellite drift winds



- GPS radio occultation



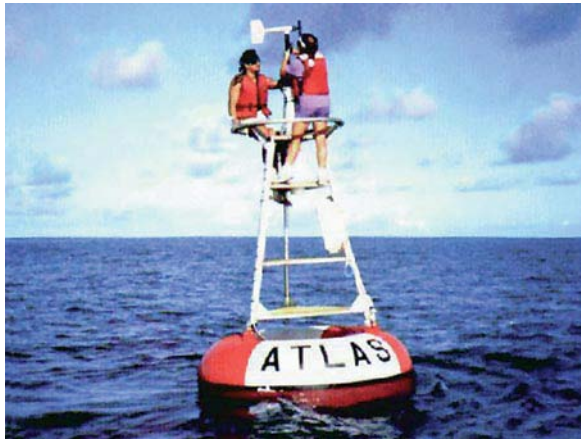
Observations of moisture and pressure are also available.

Bias corrected radiance observations will be available by late 2012.

World Ocean Database T,S 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.

Other observation types (for polar)

Already have interfaces in DART (but need testing in CESM assimilation):

- ✓ AOD aerosols
- ✓ MODIS Cloud quantities
- ✓ TES chem constituents (?)
- ✓ AIRS CO2
- ✓ GTS buoy, ship, SYNOP,AMDAR wind,T,ps
- ✓ Radar reflectivity, precip fall speed (radars at the poles?)
- ✓ Altimeter
- ✓ Dewpoint

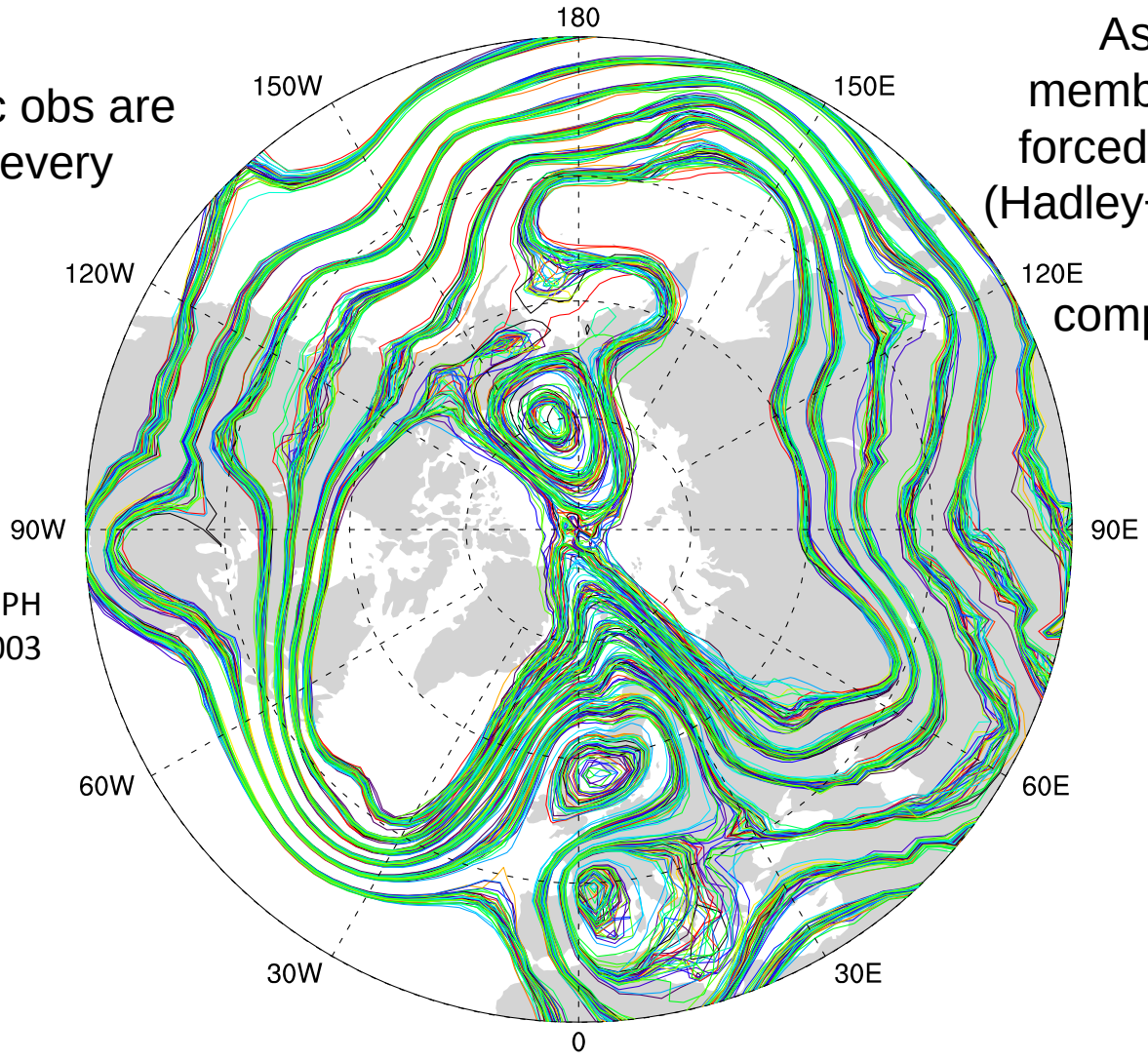
Need work to use in CAM assimilations:

- o METAR 10m winds
- o Surface observations; surface pressure (whose surface?)
- o Radiances from satellites
- o …?

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 2º FV CAM forced by a single ocean (Hadley+ NCEP-OI2) and produces a very competitive reanalysis.

Forecasts from DART/CAM analysis show the influence of a real sea ice anomaly on model surface pressure, temperatures, and low clouds.

Forecasts and figure from Jen Kay.

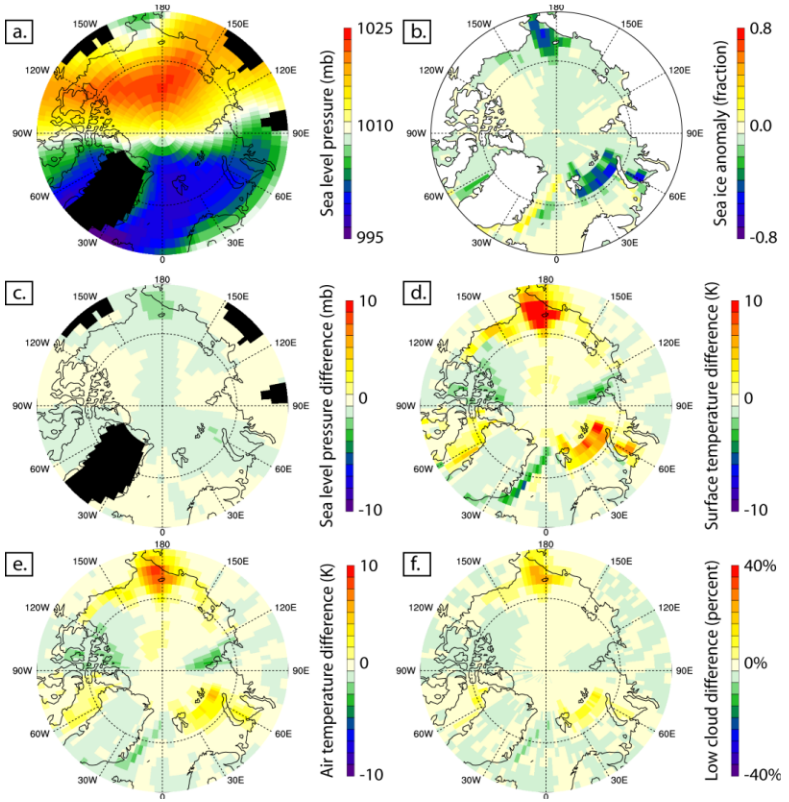
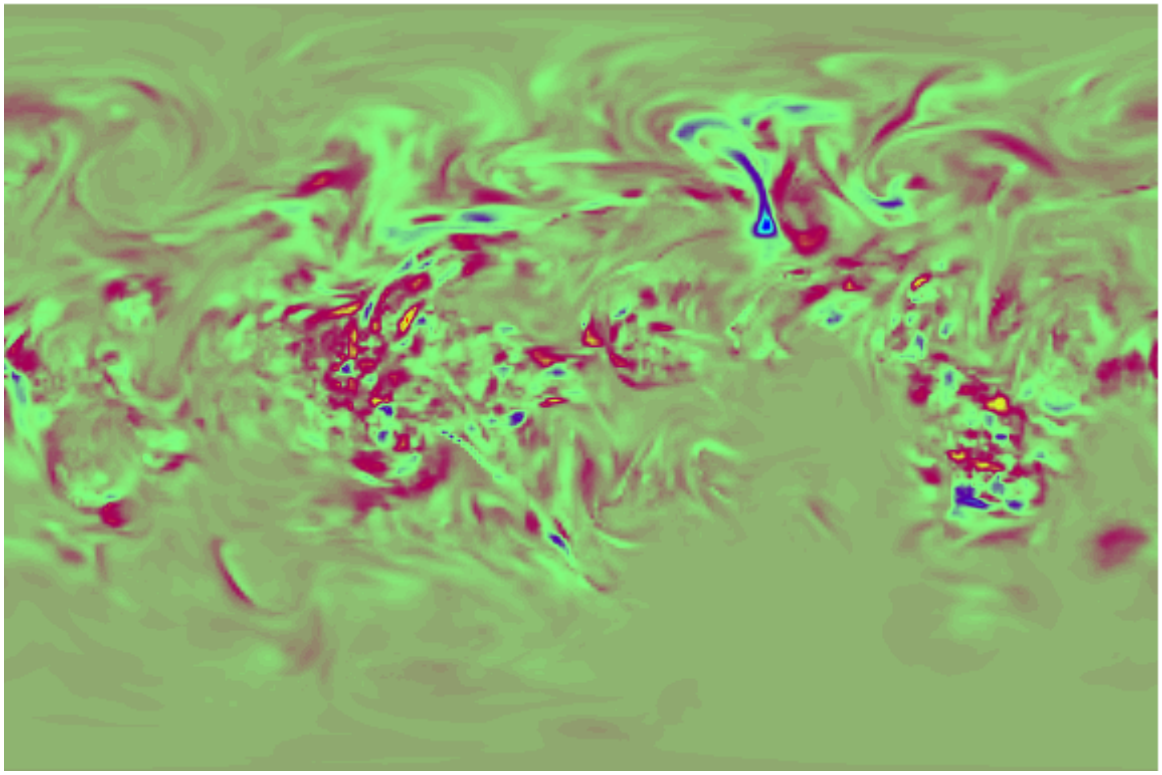


Figure x. Influence of December 2007 sea ice anomaly on December 2007 Arctic atmosphere in CAM-DART forecasts (Dec07-Dec07clim 24 hour forecasts): a) December 2007 sea level pressure, b) December 2007 sea ice anomaly, c) influence of sea ice anomaly on Arctic sea level pressure (mb), d) influence of sea ice anomaly on Arctic surface temperature (T), e) influence of sea ice anomaly on Arctic air temperature in the lowest model layer (T), and f) influence of sea ice anomaly on Arctic low cloud occurrence (percent).

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

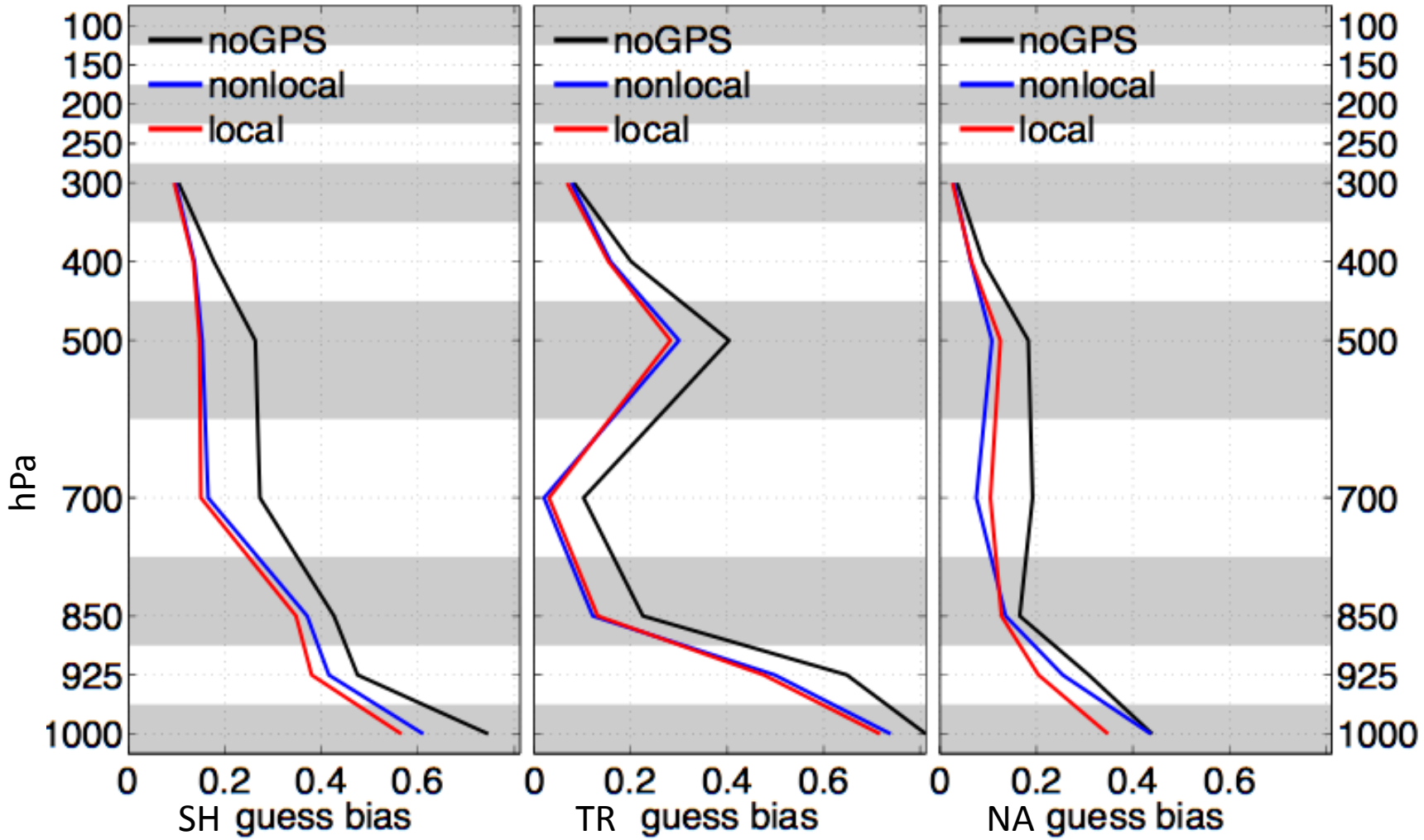


- ~200 hPa
- -.06 < ΔQ < .05 g/kg

(-.05 < Q < .10 g/kg)

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

December 2006



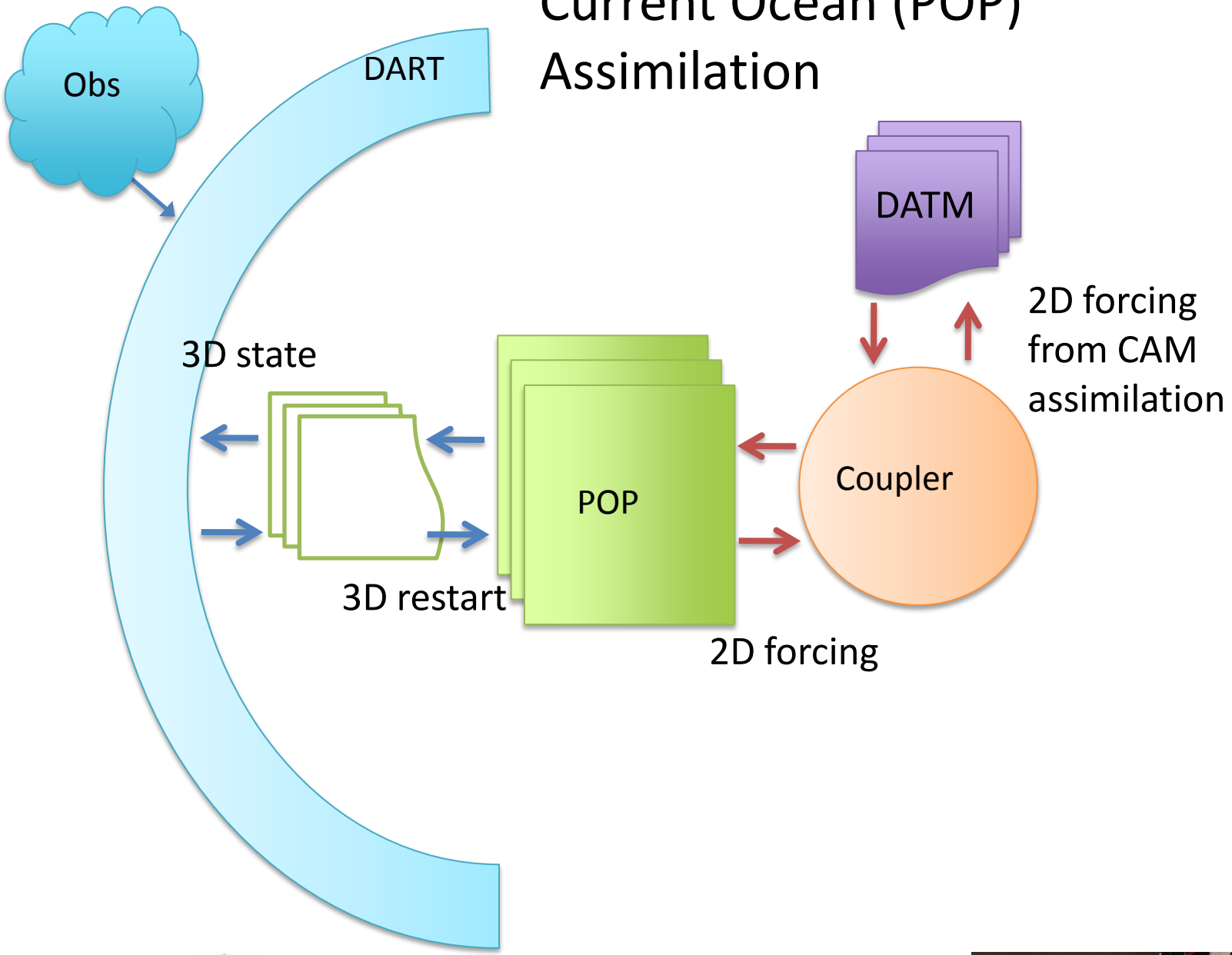
Ensemble DA Sensitivity of Cyclone Central Pressure to Mean Sea Level Pressure (left) and 300 hPa geopotential height (right)

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

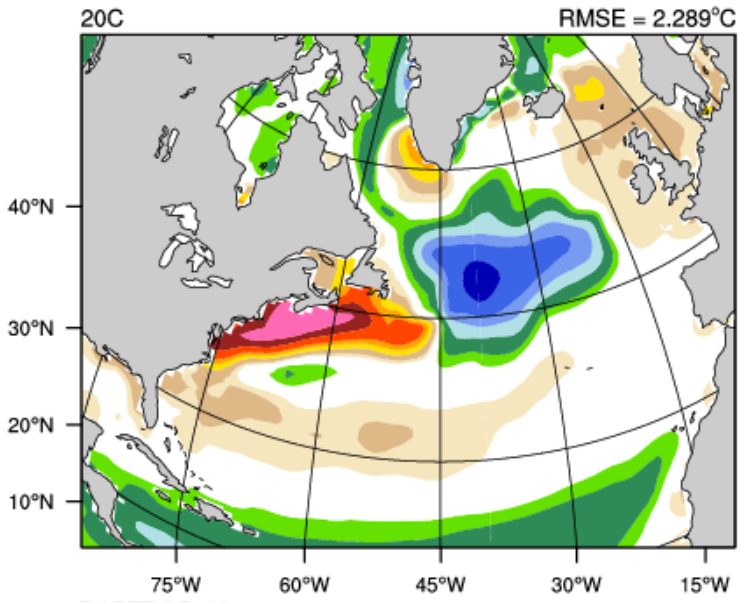
(Edmund Chang, et al., submitted to Monthly Weather Review 2011)

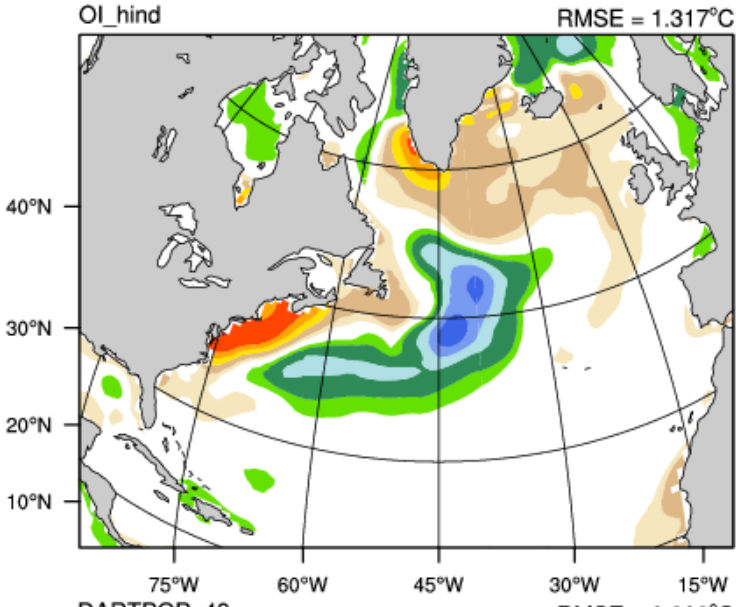
Current Ocean (POP) Assimilation

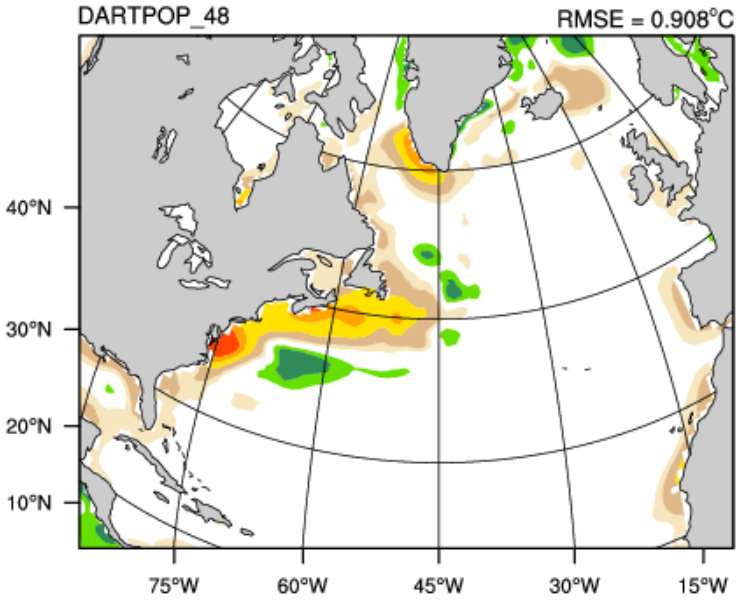


Physical Space: 1998/1999 SST Anomaly from HadOI-SST

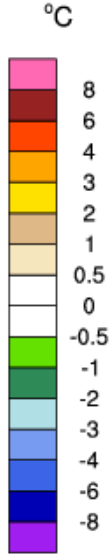
Coupled Free Run







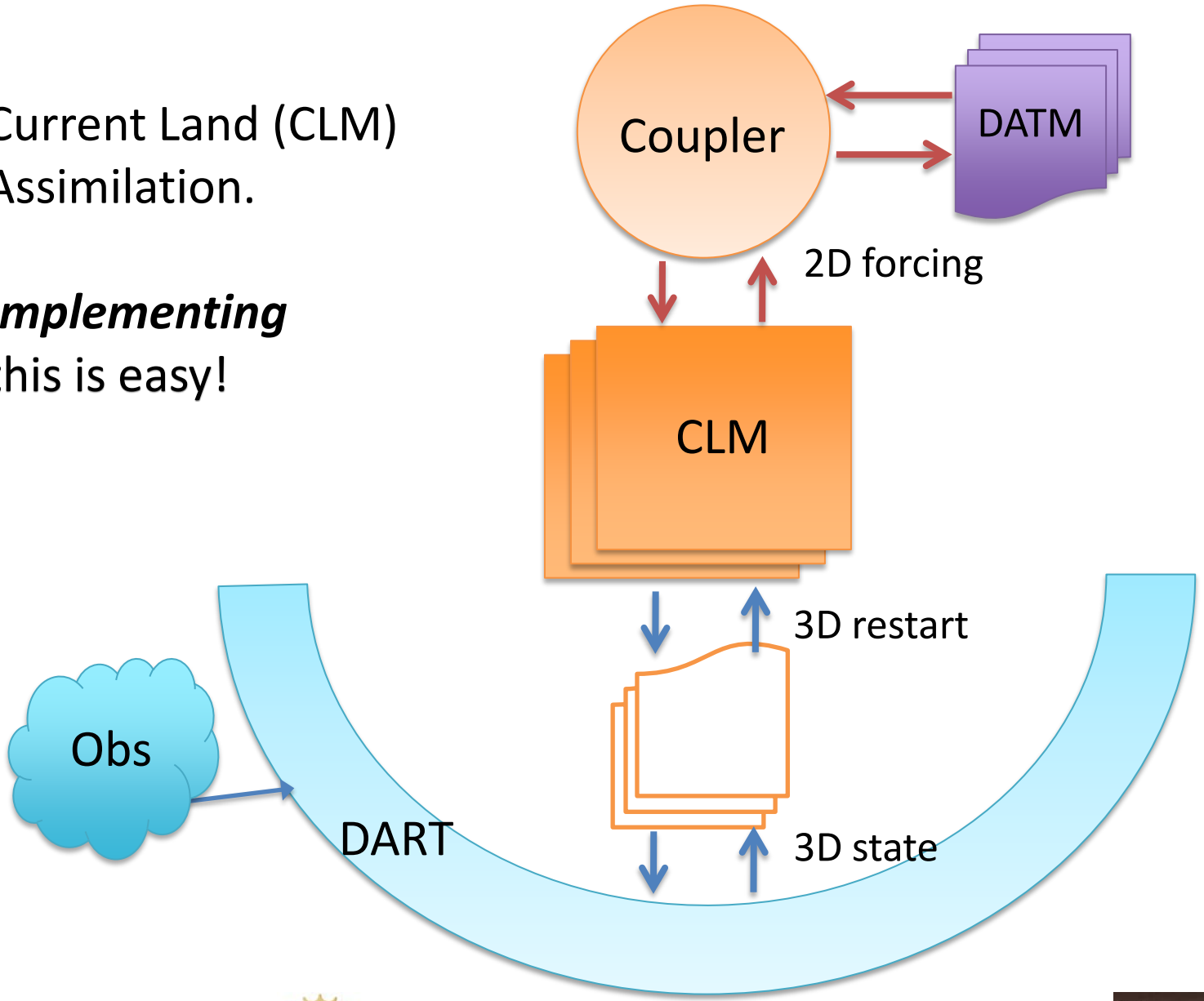
48 POP 48 CAM



POP forced by observed atmosphere (hindcast)

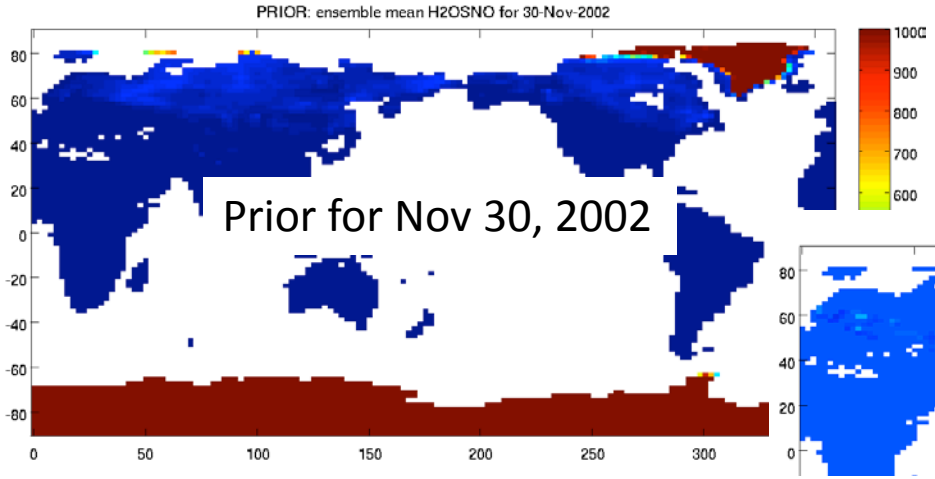
Current Land (CLM) Assimilation.

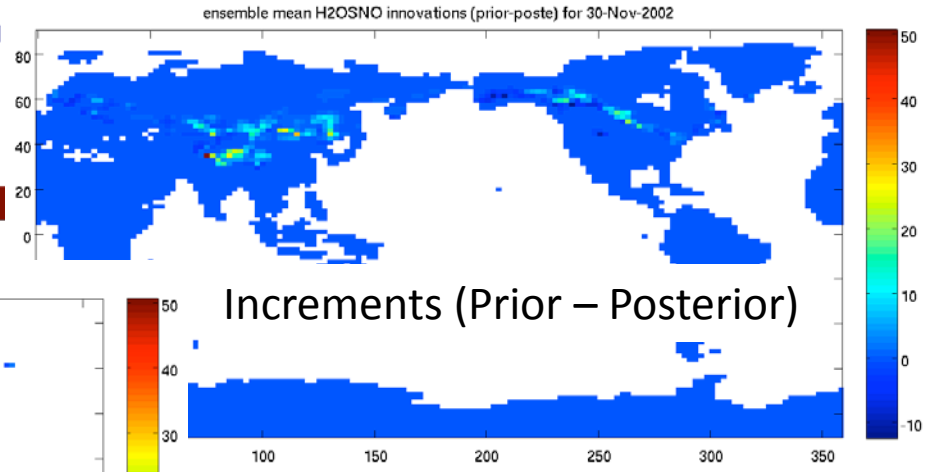
Implementing
this is easy!

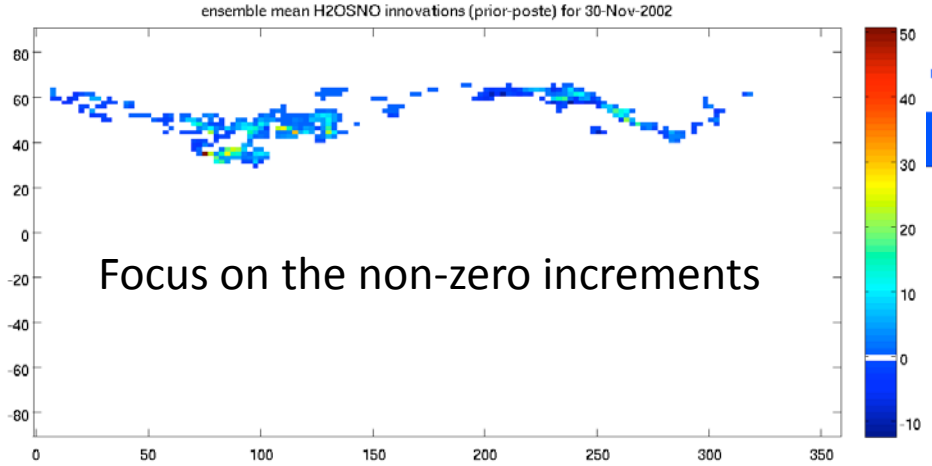


2D forcing from CAM assimilation

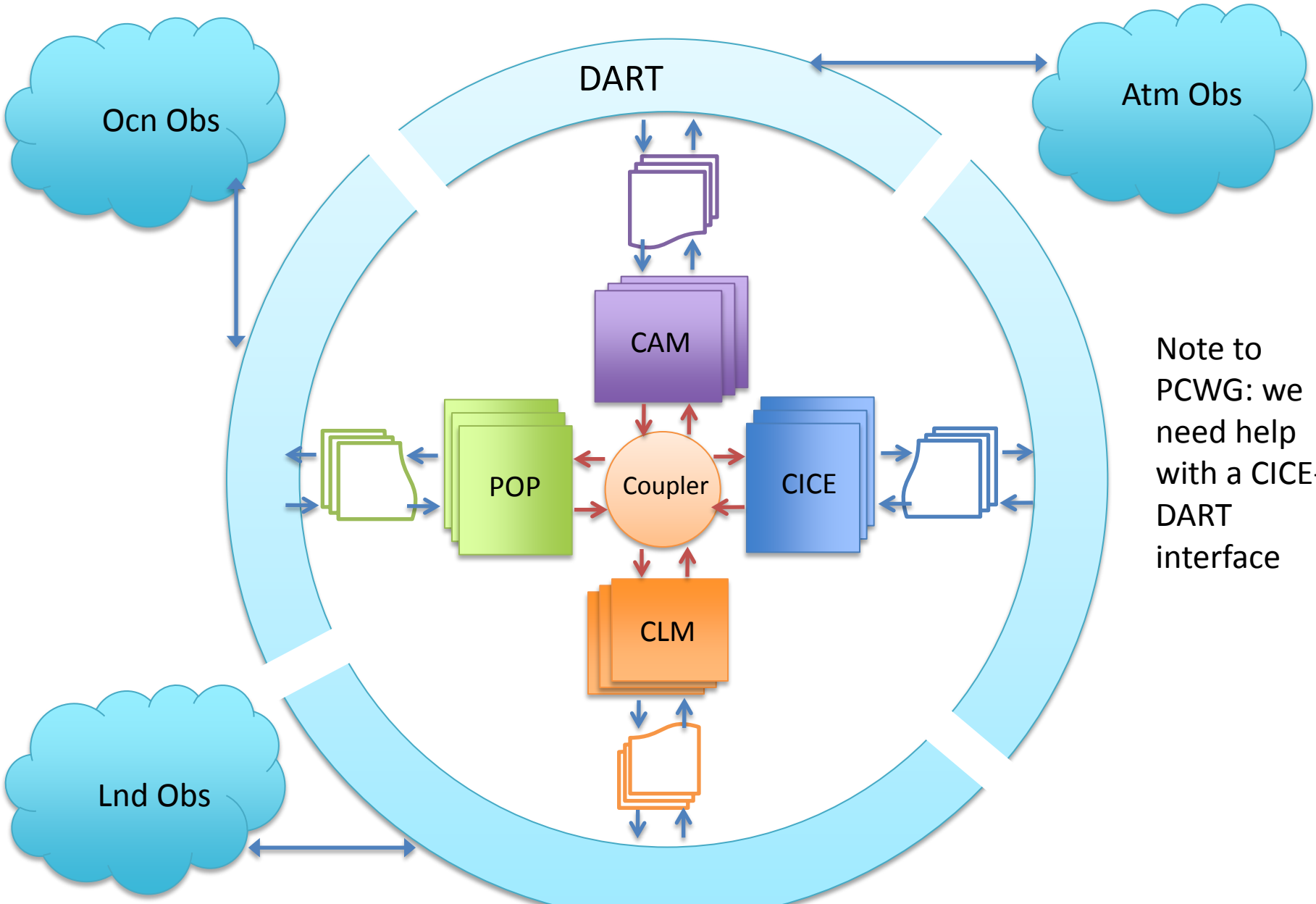
If we restrict ourselves to the simple cases … here is the early result of an assimilation of MODIS snowcover fraction on total snow water equivalent in CLM.





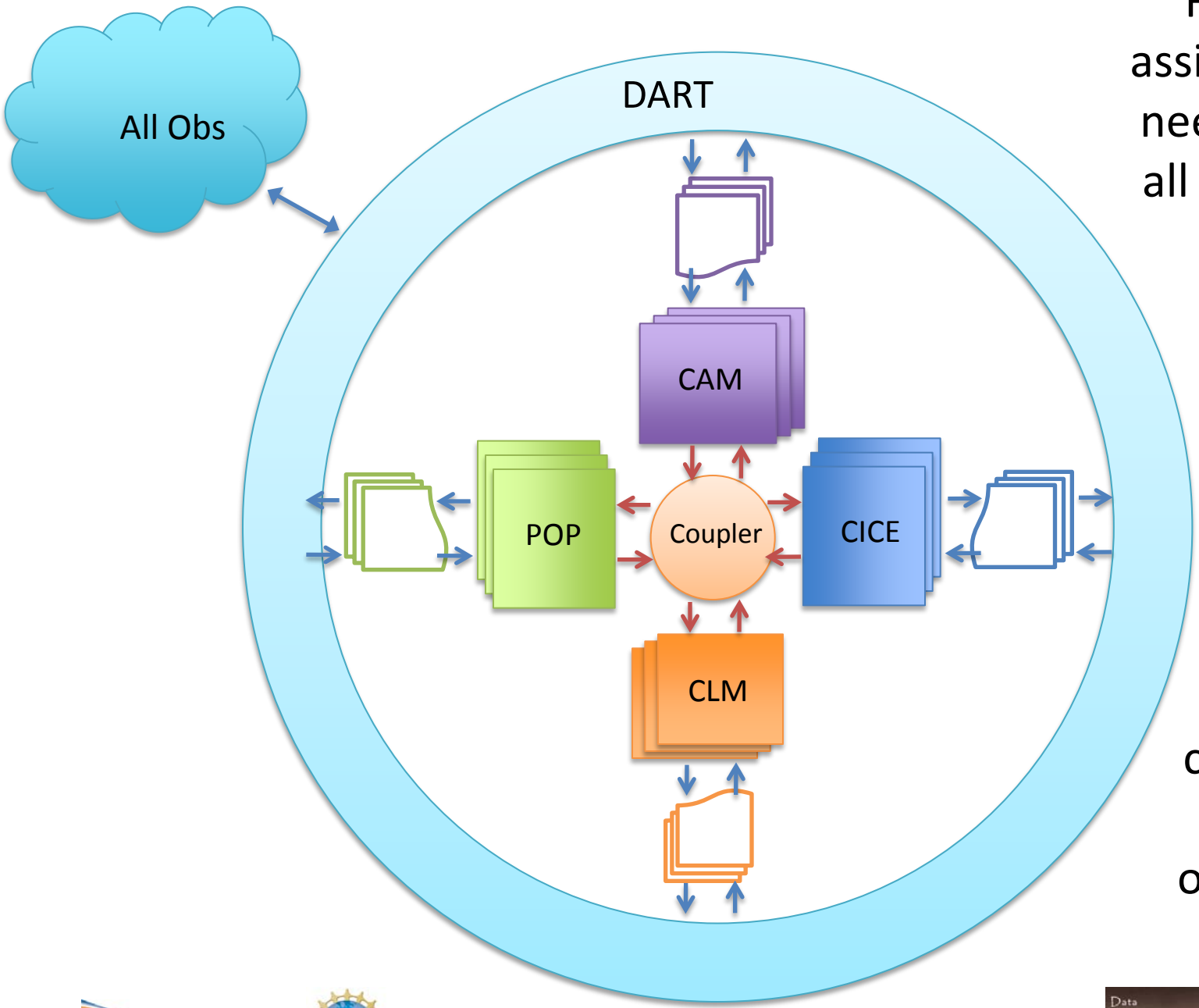


The model state is changing in reasonable places, by reasonable amounts. At this point, that's all we're looking for.



Note to PCWG: we need help with a CICE-DART interface

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



Each component corrected by all kinds of observations

Plans and Challenges

- Build fully coupled earth system model ensemble assimilation system.
- CICE: next component
- Ice Sheet Model: Challenging and interesting!
- Methods for dealing with land surface variables.
- Methods for dealing with strongly biased models.

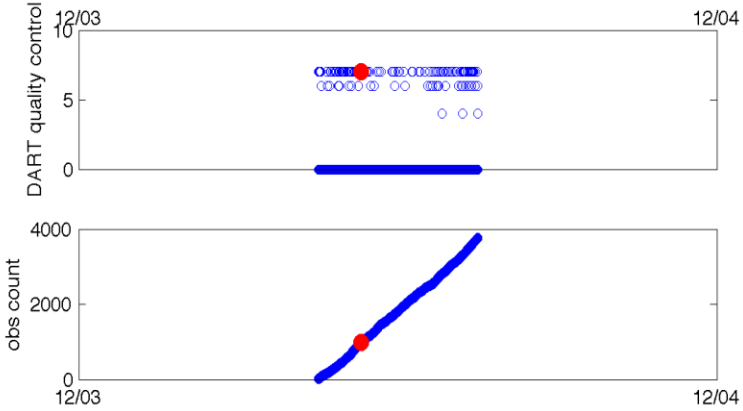
Summary

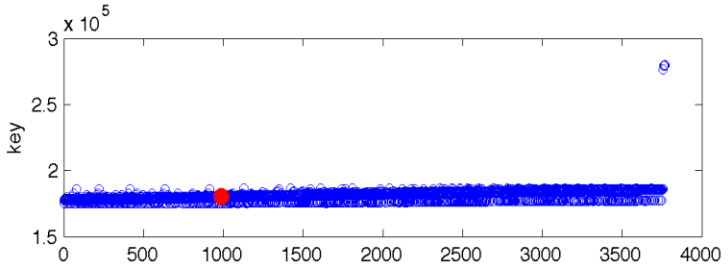
CESM components+DART ensemble DA exists for:

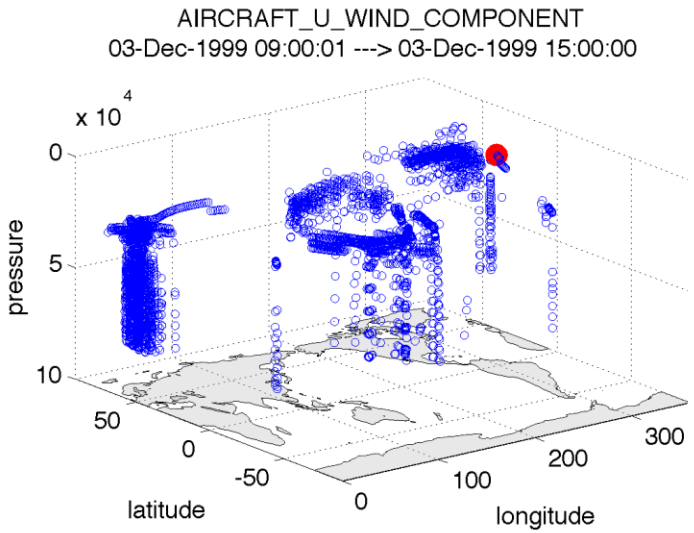
- CAM: Multiyear ensemble reanalysis available,
- POP: Ensemble analyses used for decadal prediction initial conditions,
- CLM: Ensemble snow cover analyses and leaf area index research.

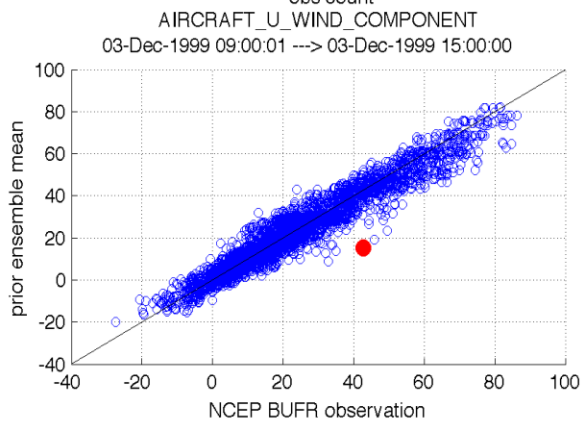
Extras

slide held in reserve









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"?

Crusty Snow?

Slushy Snow?

Old Snow?

Dirty Snow?

Early Season Snow?

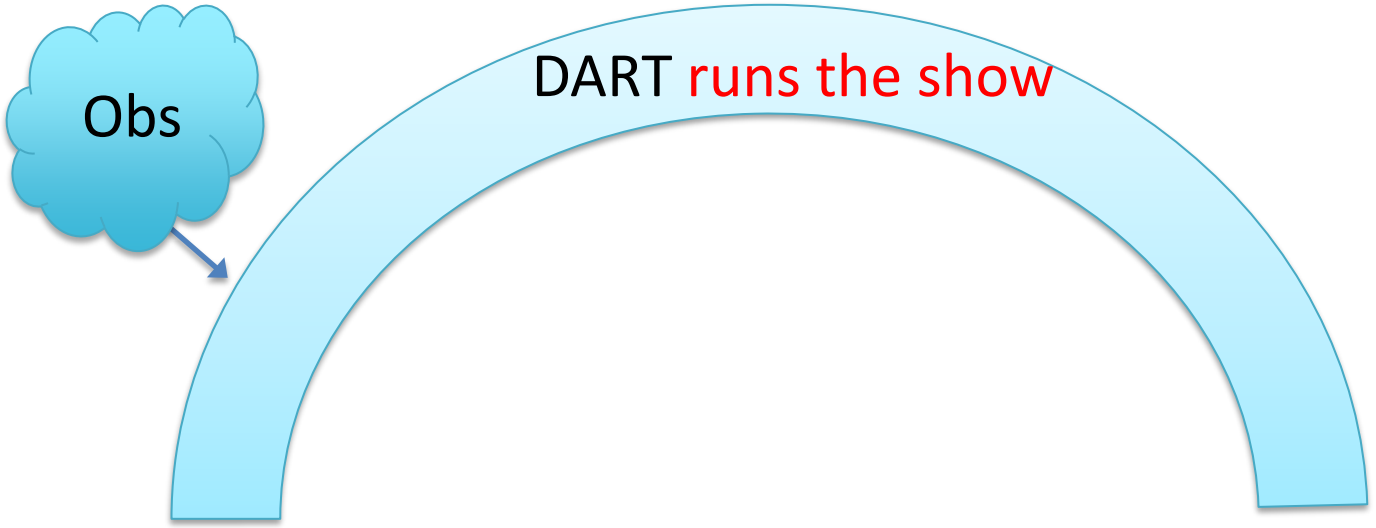
Packed Snow?

Snow Density?

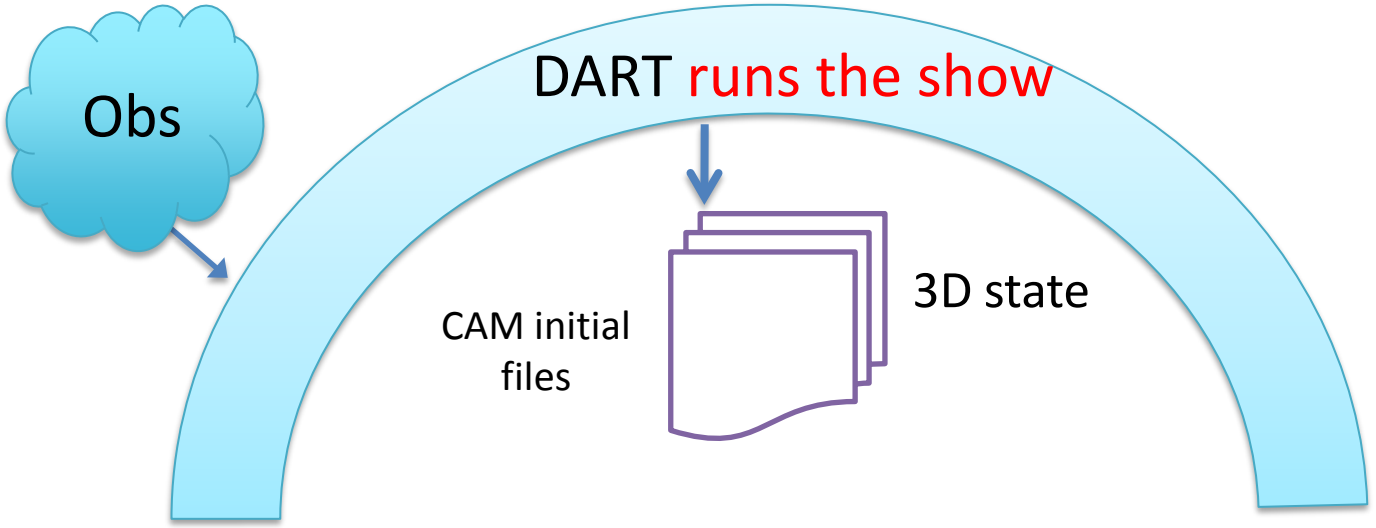
Snow Albedo?

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!

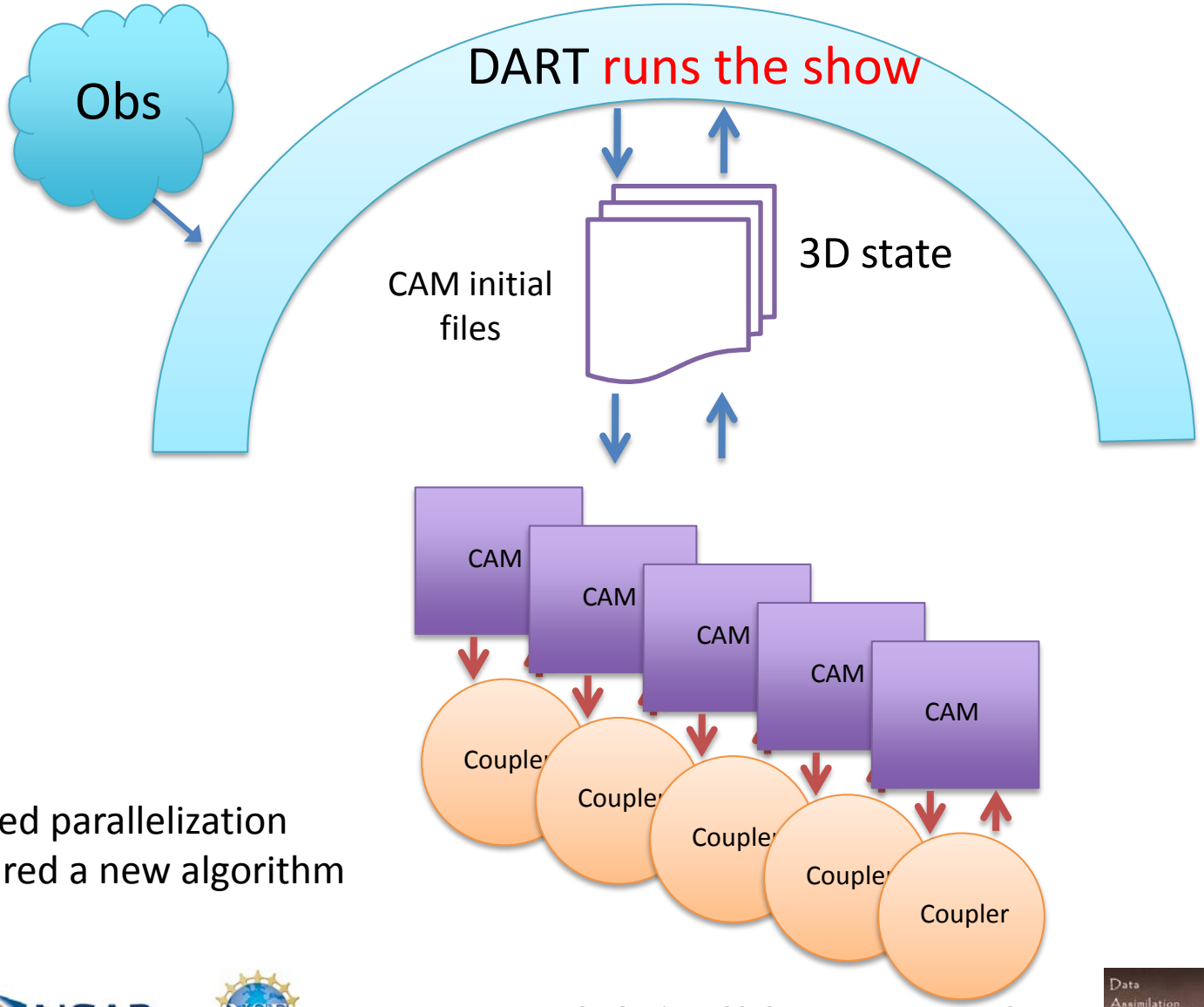
Past CAM Assimilation



Past CAM Assimilation



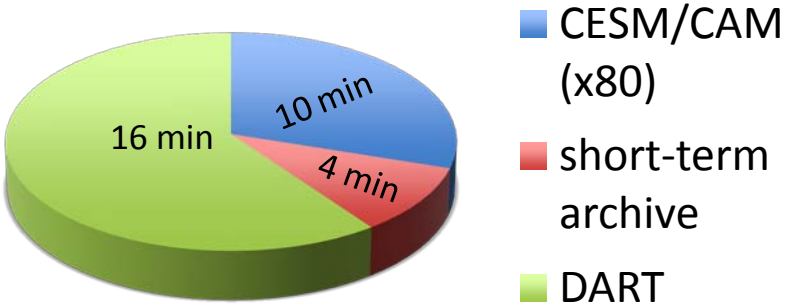
Past CAM Assimilation



Limited parallelization required a new algorithm

Cost/Efficiency

1-deg CAM4+DART 6 hour Assimilation



- 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)

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!