

DART and Land Data Assimilation



Tim Hoar: *National Center for Atmospheric Research*

with a whole lot of help from:

Jeff Anderson, Nancy Collins, Kevin Raeder, Bill Sacks: *NCAR*

Yongfei Zhang: *University of Texas Austin*

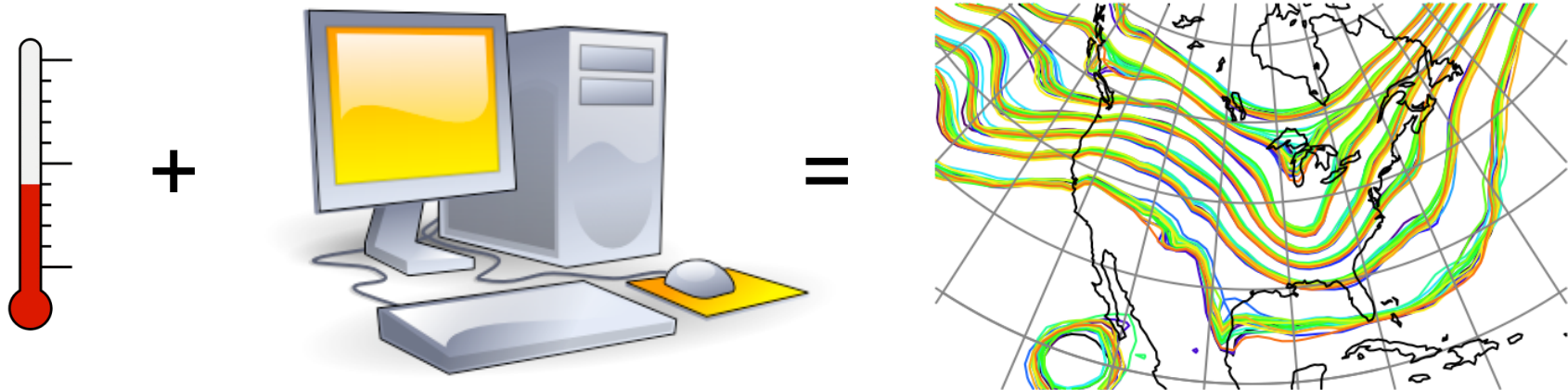
Andrew Fox: *National Ecological Observatory Network (NEON)*

Rafael Rosolem: *University of Arizona*



What is Data Assimilation?

Observations combined with a Model forecast...

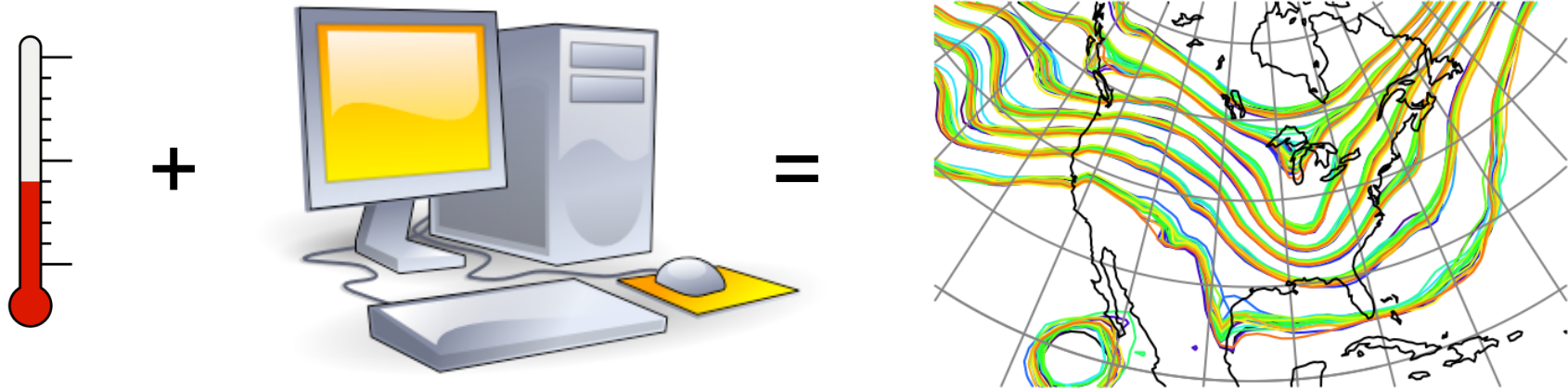


... to produce an analysis.

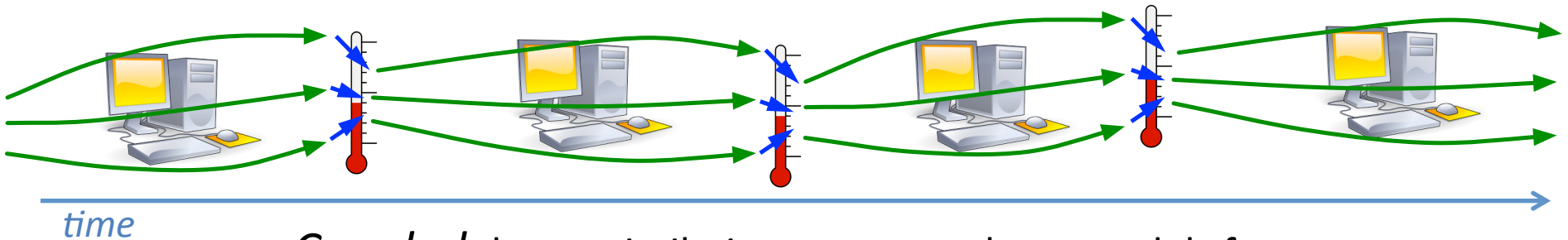
Overview article of DART:

Anderson, Jeffrey, T. Hoar, K. Raeder, H. Liu, N. Collins, R. Torn, A. Arellano, 2009:
The Data Assimilation Research Testbed: A Community Facility.

Bull. Amer. Meteor. Soc., **90**, 1283–1296. [doi:10.1175/2009BAMS2618.1](https://doi.org/10.1175/2009BAMS2618.1)



We want to assimilate over and over to steadily make the model states more consistent with the observations.



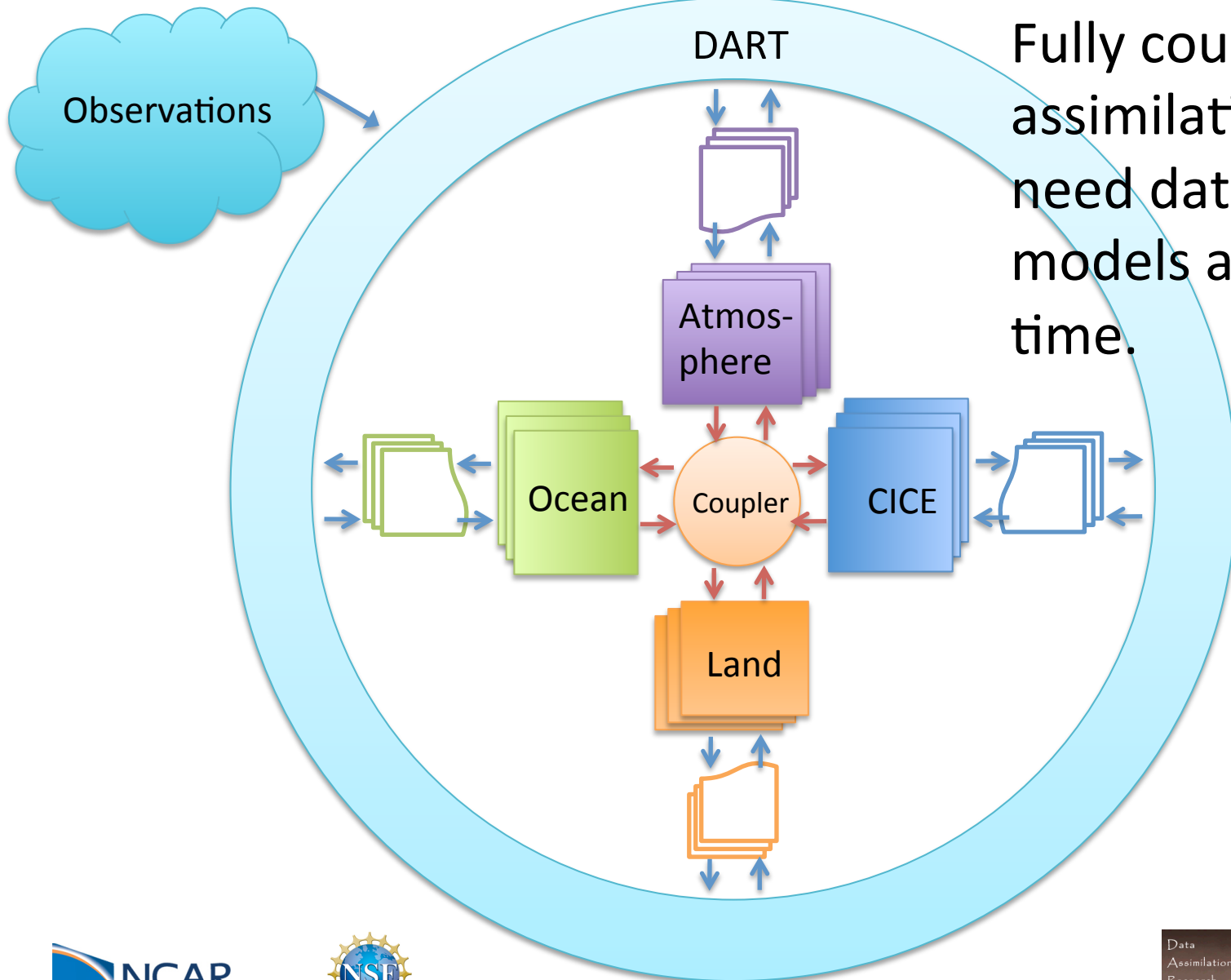
Coupled data assimilation means we have models for atmosphere and ocean, or atmosphere and land, or all three, or ...



A short list of models that can assimilate with DART:

- CAM: Community Atmosphere Model
- POP: Parallel Ocean Program
- WRF: Weather Research and Forecasting Model
- AM2: GFDL Atmospheric Model
- COAMPS: Coupled Atmosphere/Ocean Mesoscale Prediction System
- CLM: Community Land Model
- NOAH: Land Surface Model
- ... many more



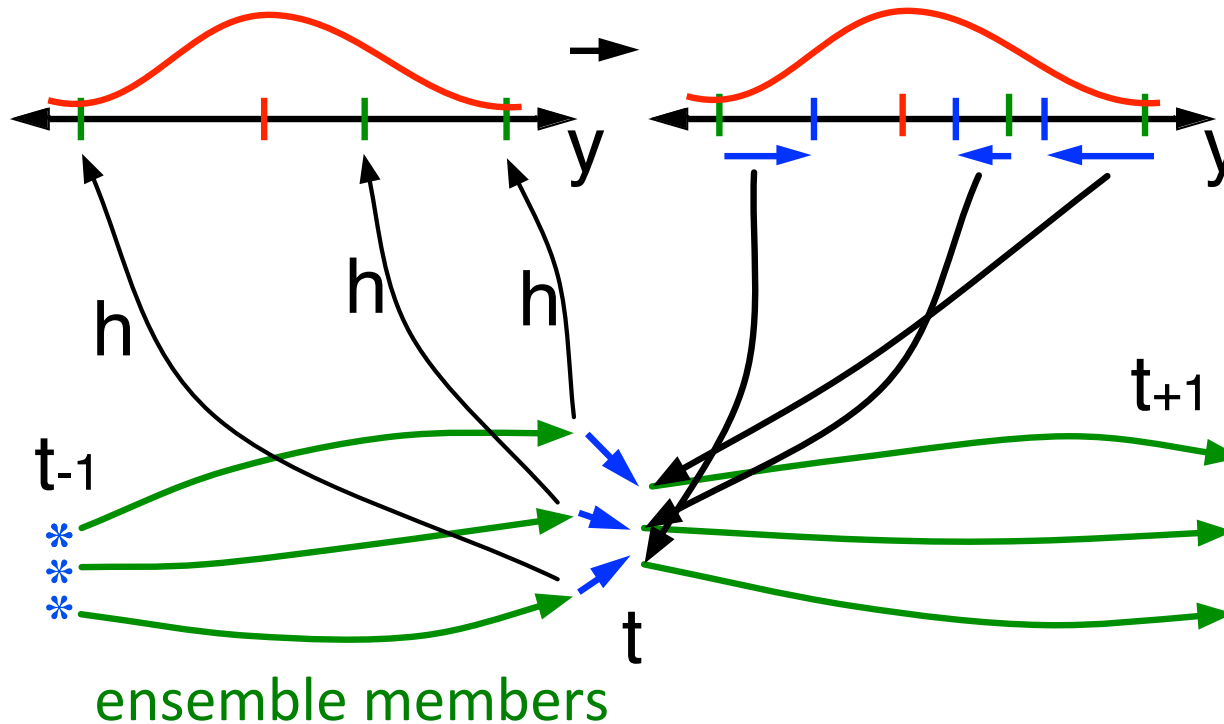


Fully coupled assimilation will need data from all models at the same time.



A generic ensemble filter system like DART needs:

1. A way to make model forecasts.
2. A way to estimate what the observation would be – given the model state. This is the forward observation operator – h .



The **increments** are regressed onto as many **state variables** as you like. If there is a correlation, the state gets adjusted in the restart file.

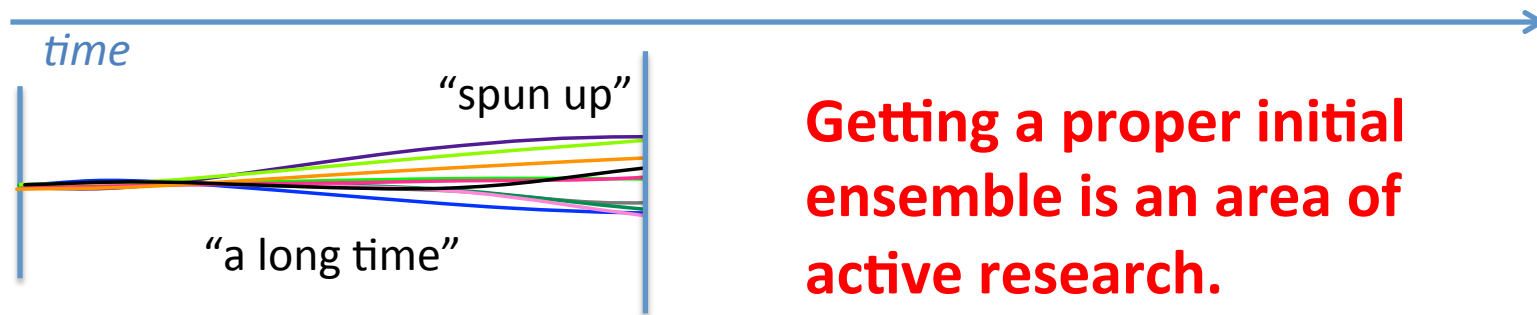




- One unbreakable rule: “**Do Not Invade** the model code.”
- Unique routines communicate between each model and DART to provide this separation.
- I want to use COSMOS observations with “all” of our land models. This means our observation operator must be pretty generic.
- The land models are an area where observations are needed to help constrain the model states – so we can learn about and improve the models.



Creating the initial ensemble of ...

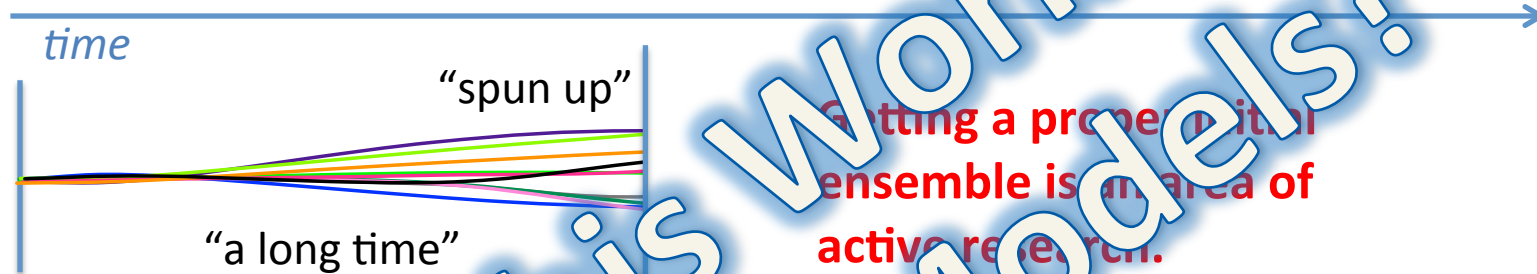


1. Replicate what we have N times.
2. Use a unique (and different!) *realistic* forcing for each.
3. Run them forward for "a long time".

We have tools we are using to explore how much spread we NEED to capture the uncertainty in the system.



Creating the initial ensemble of ...



1. Replicate what we have N times.
2. Use a unique (and different) *realistic* forcing for each.
3. Run them forward for "a long time".

We have tools we are using to explore how much spread we NEED to capture the uncertainty in the system.

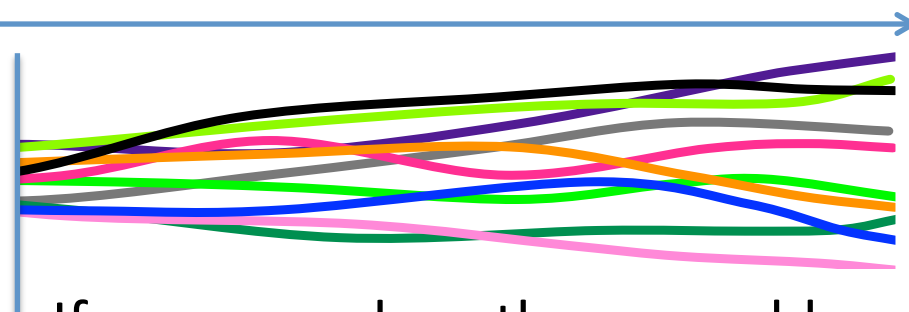


The ensemble advantage.

You can represent uncertainty.

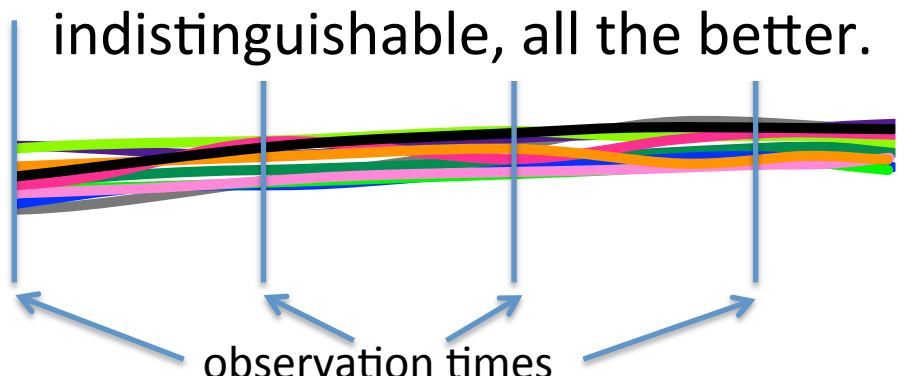
time

The ensemble spread frequently grows in a free run of a dispersive model.



If we can reduce the ensemble spread and still be indistinguishable, all the better.

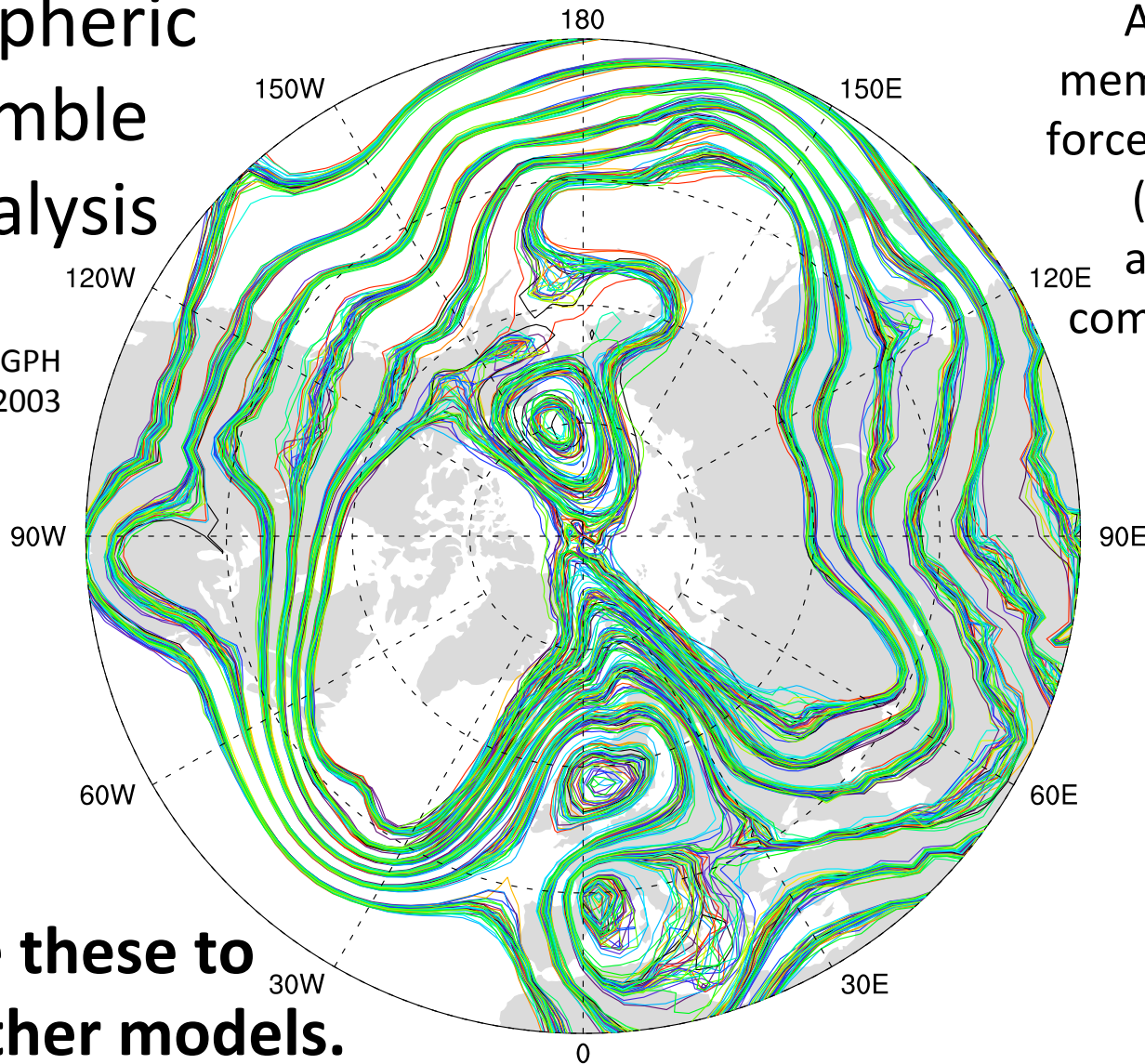
With a good assimilation the ensemble is indistinguishable from the true model state in any meaningful way.





Atmospheric Ensemble Reanalysis

500 hPa GPH
Feb 17 2003



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

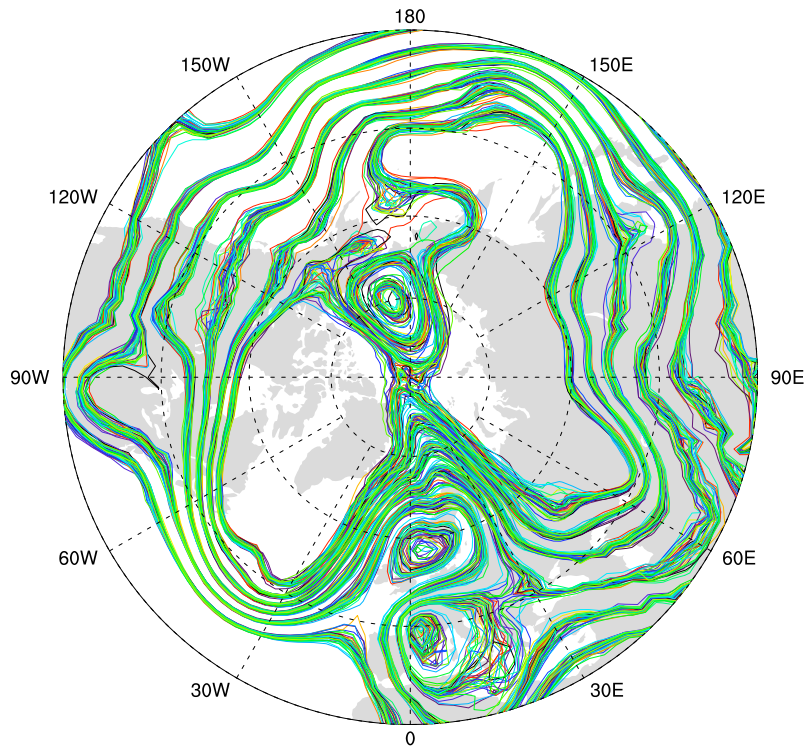
O(1 million) atmospheric obs are assimilated every day.

1998-2010
4x daily is available.

Can use these to force other models.



Pros and Cons

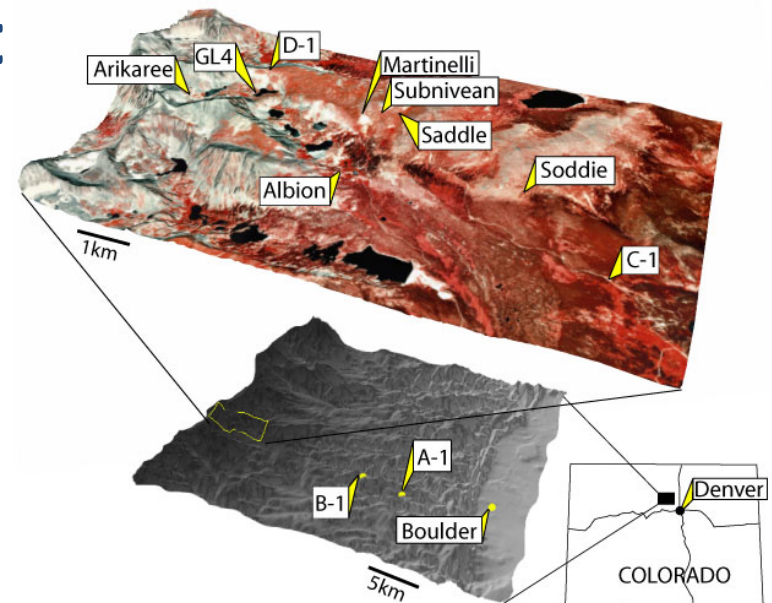


- **80 realizations/members**
- **Model states are self-consistent**
- **Model states consistent with observations**
- **Available every 6 hours**
- Relatively low spatial resolution has implications for regional applications.
- Suboptimal precipitation characteristics.
- Available every 6 hours
 - higher frequency available if needed.
- Only have 12 years ... enough?

Since Rafael is going to be showing results of NOAH, I'll explore some of these (and other) issues with results from CLM – one for a flux tower, one for global snow data assimilation.

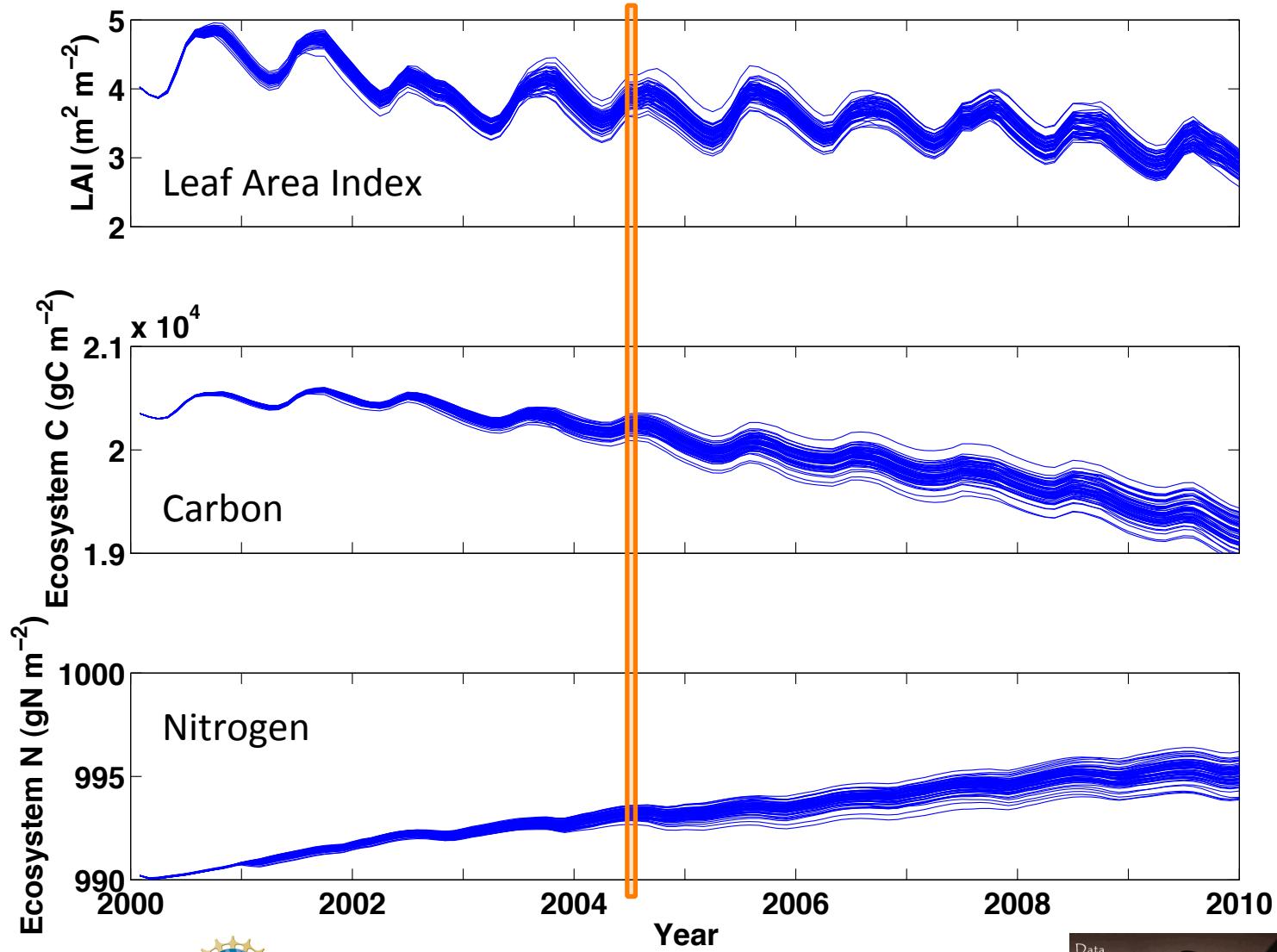
In collaboration with Andy Fox (NEON): An experiment at Niwot Ridge

- 9.7 km east of the Continental Divide
- C-1 is located in a Subalpine Forest
- (40° 02' 09" N; 105° 32' 09" W; 3021 m)
- Single column of Community Land Model
- 64 ensemble members of CLM
- Forcing from the DART/CAM reanalysis,
- Assimilating tower fluxes of latent heat (LE), sensible heat (H), and net ecosystem production (NEP).
- Impacts CLM variables: LEAFC, LIVEROOTC, LIVESTEMC, DEADSTEMC, LITR1C, LITR2C, SOIL1C, SOIL2C, SOILLIQ ... all of these are *unobserved*.



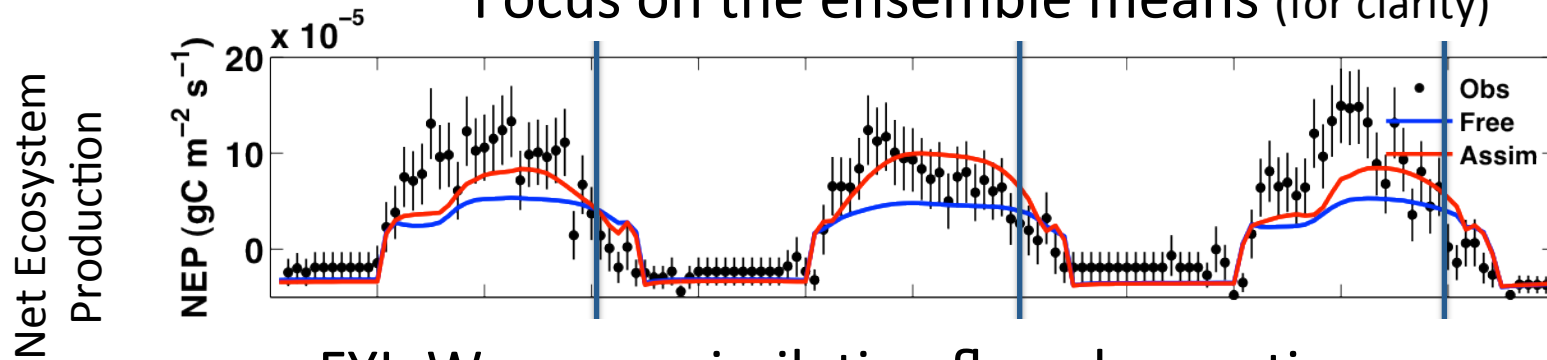


Free Run of CLM

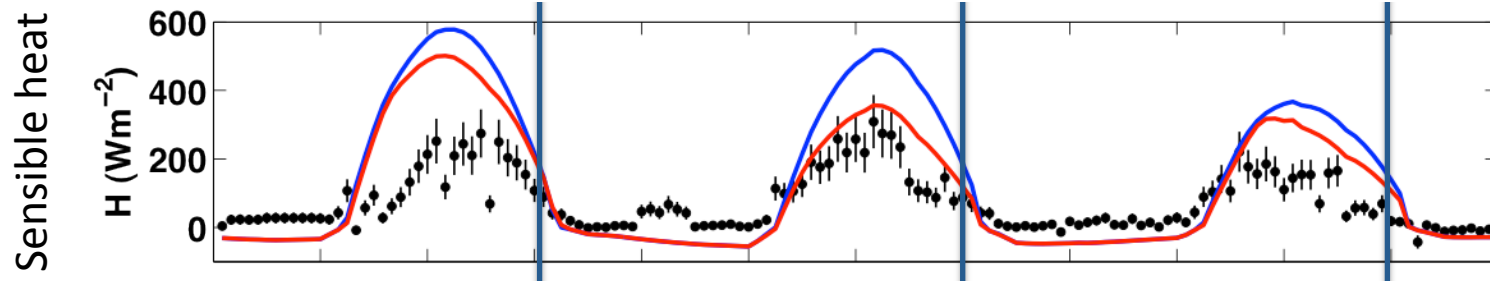




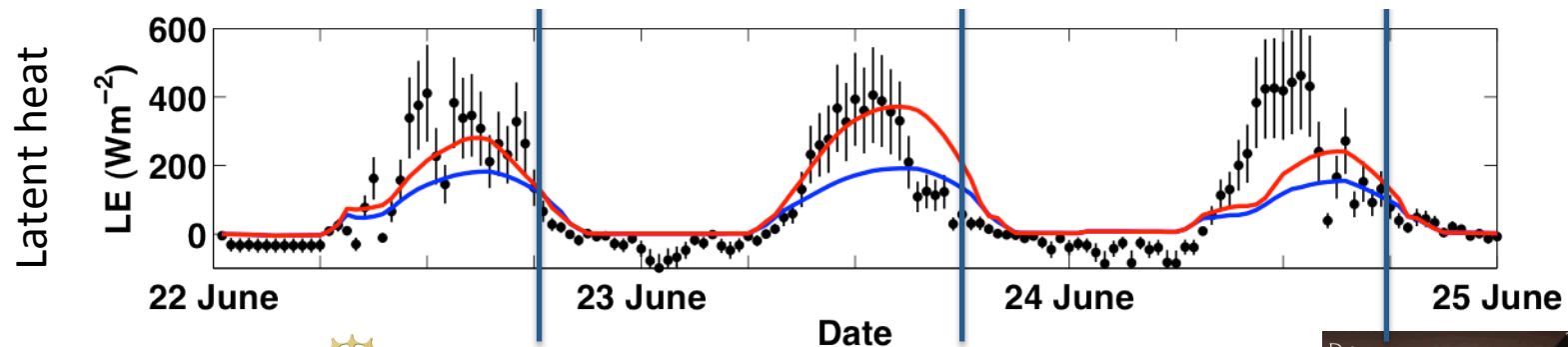
Focus on the ensemble means (for clarity)

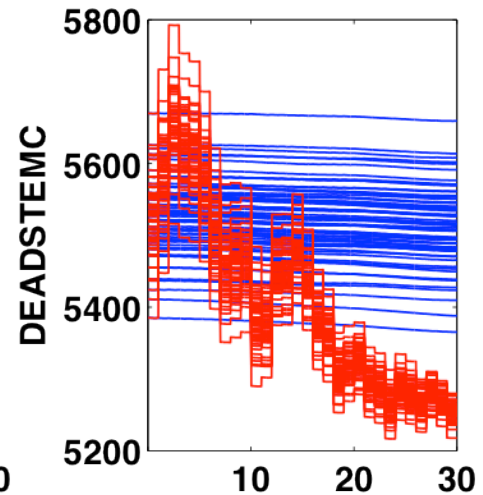
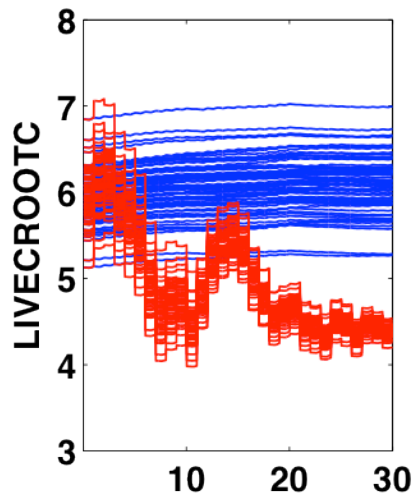
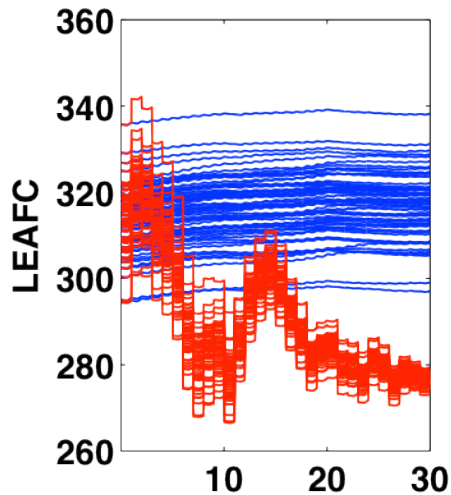


FYI: We are assimilating flux observations ...



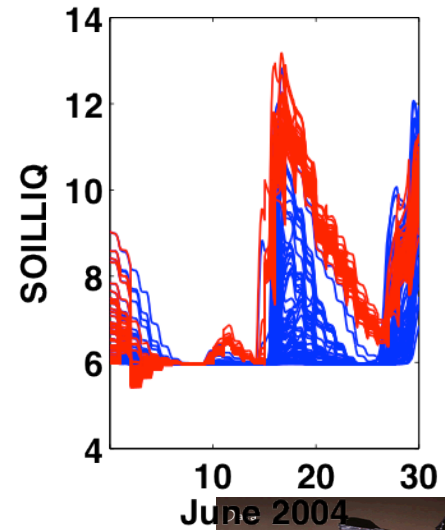
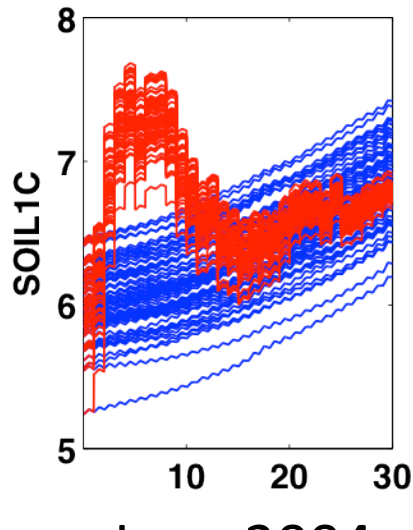
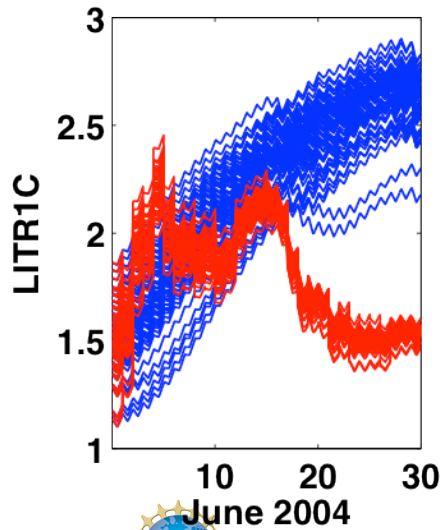
The model states are being updated at about 8PM local time.





Free Run ——— blue line
Assim ——— red line

These are all unobserved variables.

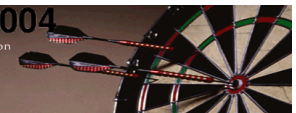


June 2004

June 2004

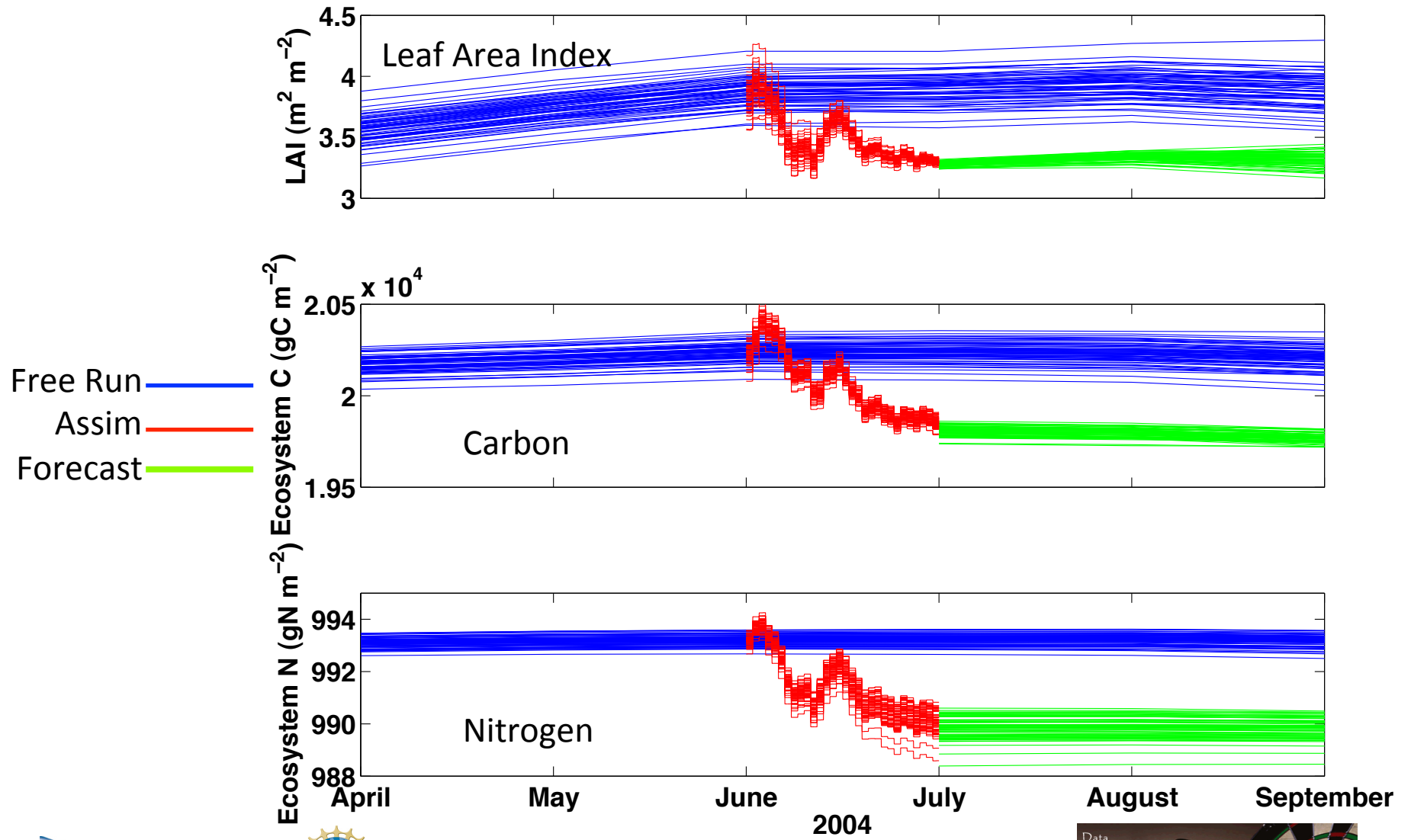
June 2004

Assimilation
Research
Testbed



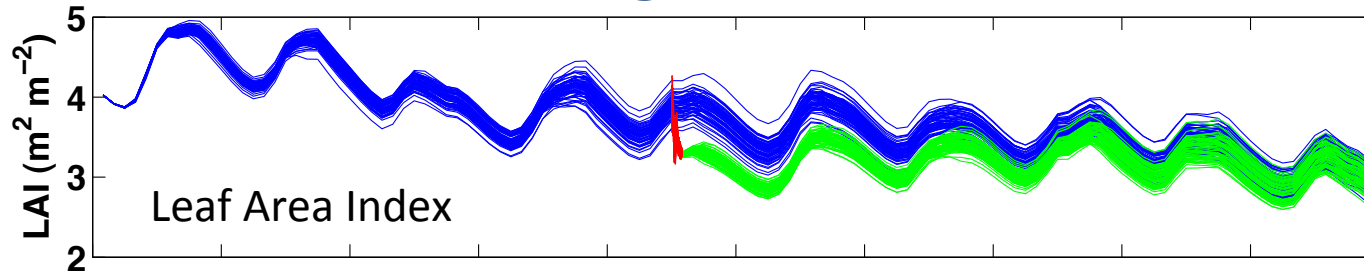


Effect on short-term forecast on unobserved variables.

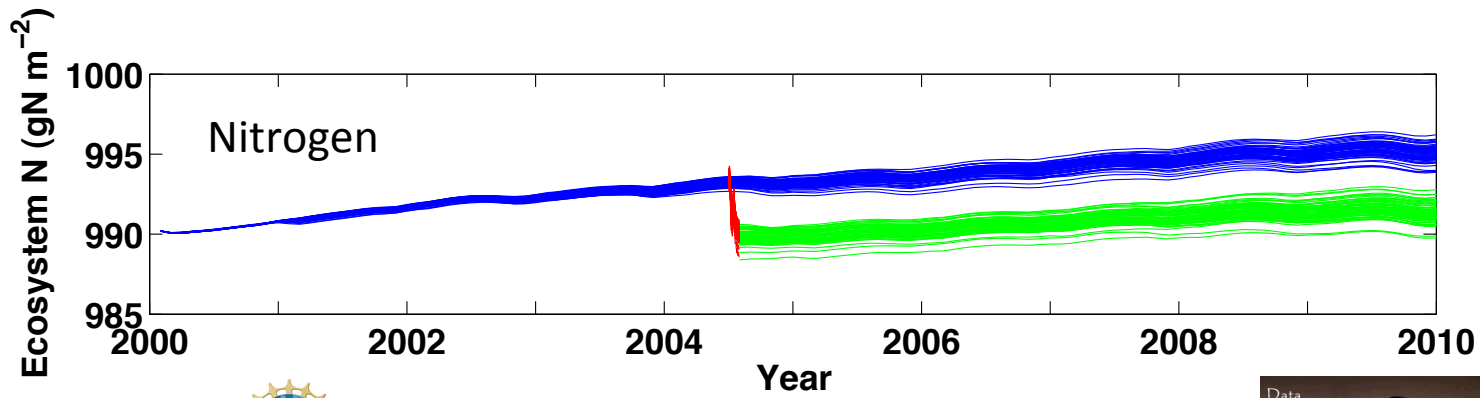
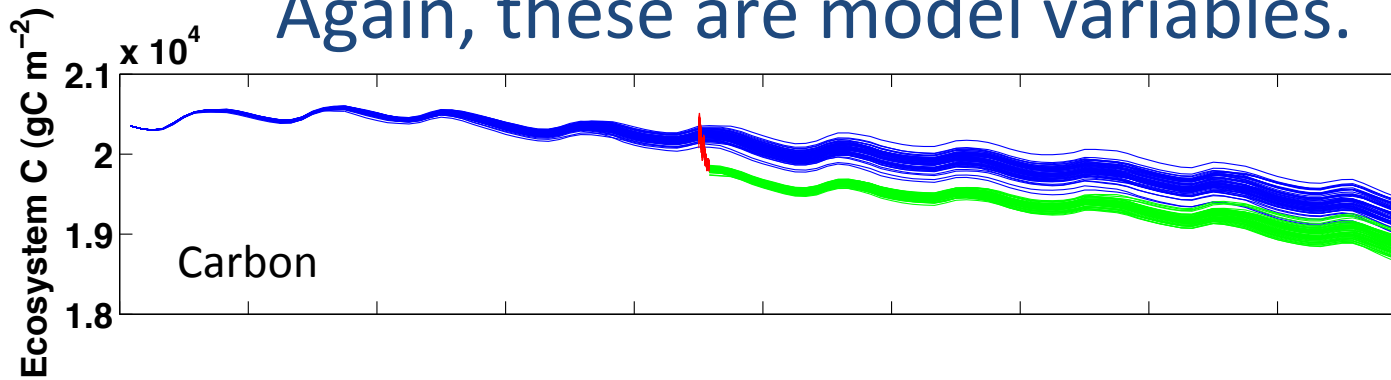


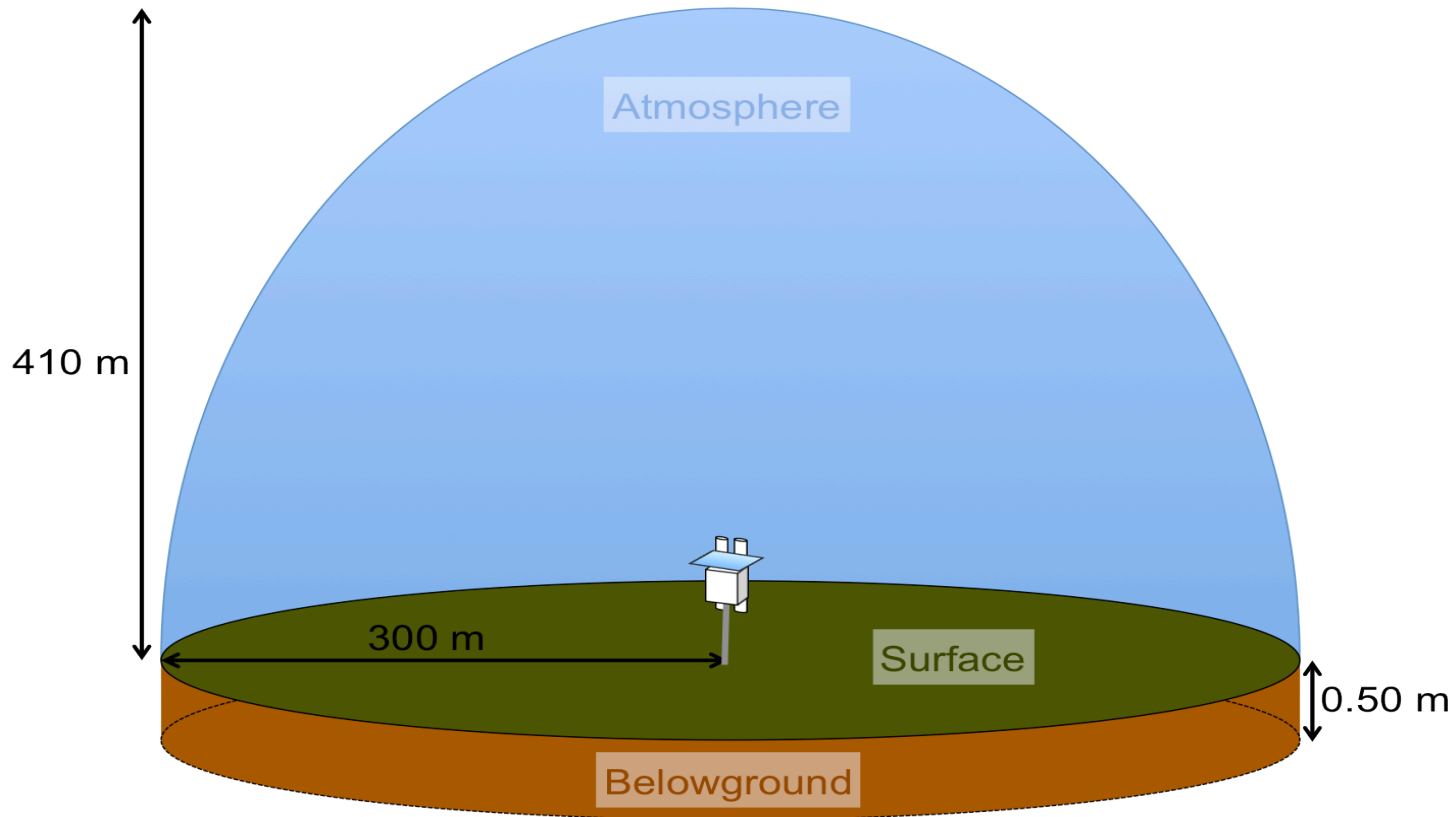


Effect on longer-term forecast



Again, these are model variables.



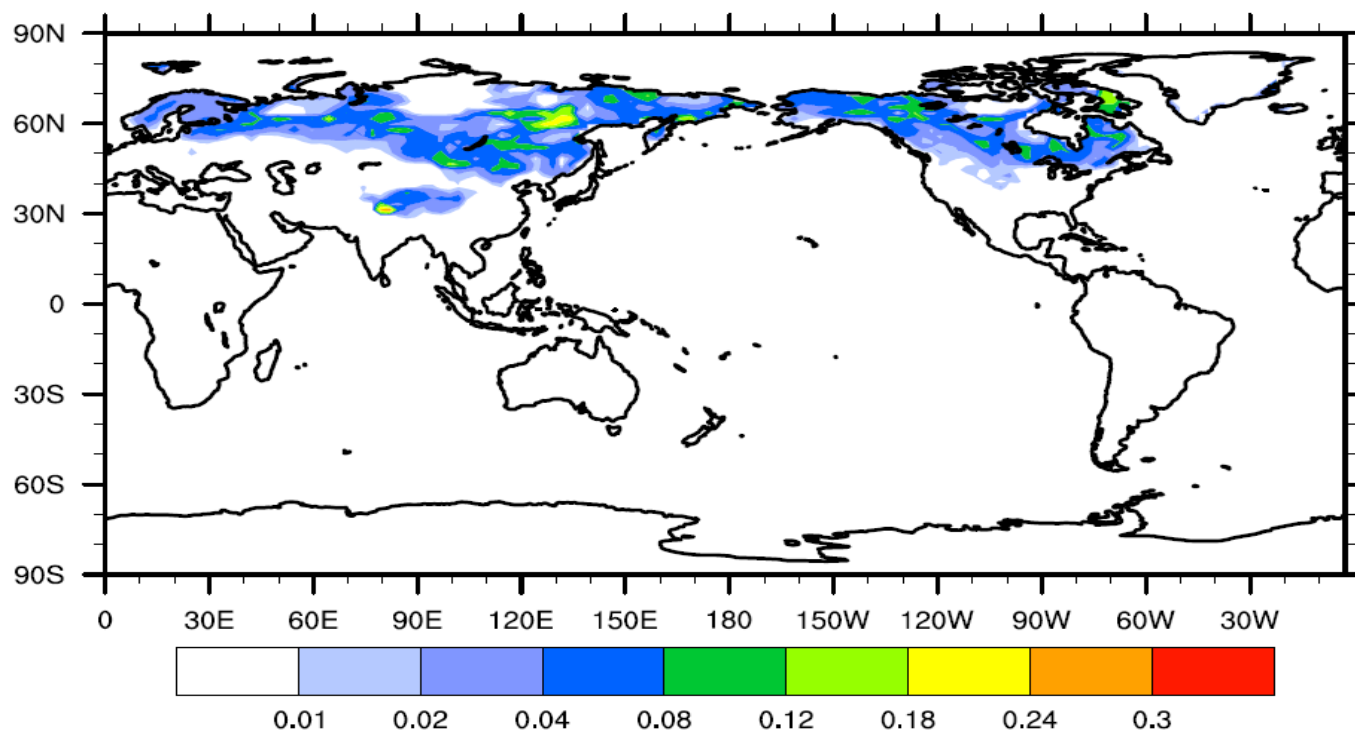




Assimilation of MODIS snow cover fraction

- 80 member ensemble for onset of NH winter, assimilate once per day
- Level 3 MODIS product – regridded to a daily 1 degree grid
- Observations can impact state variables within 200km
- CLM variable to be updated is the snow water equivalent “H2OSNO”
- **Analogous to precipitation ...**

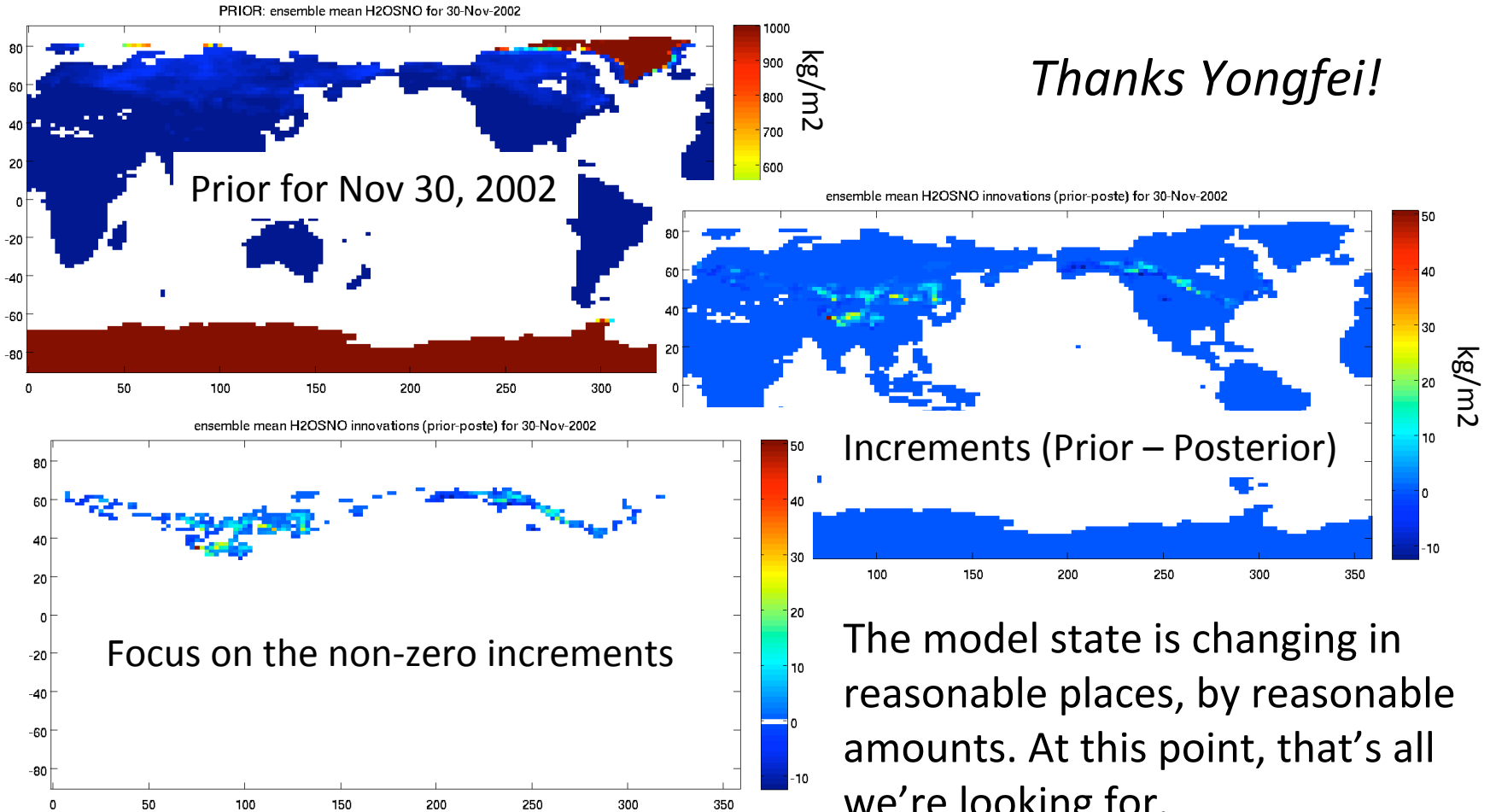
Standard deviation of the CLM snow cover fraction initial conditions for Oct. 2002





An early result: assimilation of MODIS *snowcover fraction* on total *snow water equivalent* in CLM.

Thanks Yongfei!





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

Corn Snow?

New Snow?

Sugar Snow?

Dry Snow?

Wet Snow?

“Champagne Powder”?

Slushy Snow?

Crusty Snow?

Dirty Snow?

Old 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!



As I see it, problems to be solved:

- Proper initial ensemble
- Can models tolerate new assimilated states?
- Snow ... depths, layers, characteristics, content.
 - When all ensembles have identical values the observations cannot have any effect with the current algorithms.
 - COSMOS forward observation operator for NOAH-MP, CLM ...
- Forward observation operators
 - many flux observations are over timescales that are inconvenient
 - need soil moisture from last month and now ... GRACE
 - Multisensor soil moisture assimilation?
- Bounded quantities
 - Negative SW fluxes, for example.
 - Soils dry beyond their limits.



For more information:

CAM

WRF

CLM

Data
Assimilation
Research
Testbed

POP

AM2

BGRID



COAMPS

www.image.ucar.edu/DARes/DART

NOAH

MITgcm_ocean

dart@ucar.edu

MPAS_ATM

SQG

NAAPS

MPAS_OCN

TIEGCM

COAMPS_nest

PBL_1d

NCOMMAS

PE2LYR

