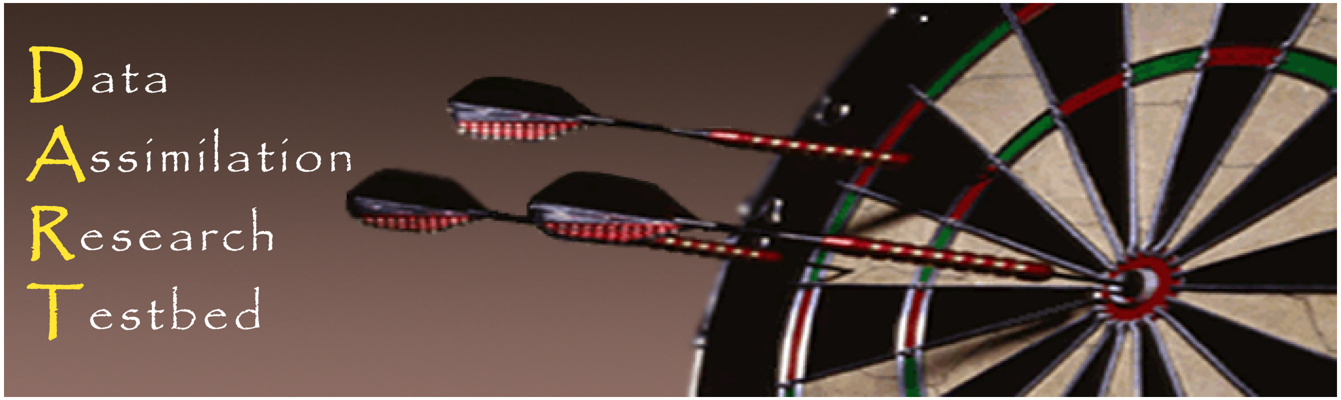
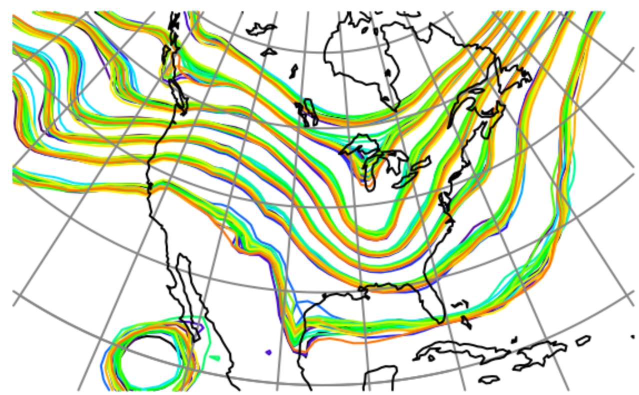


Getting to know the Data Assimilation Research Testbed - DART.

Tim Hoar, Data Assimilation Research Section, NCAR



The National Center for Atmospheric Research is sponsored by the National Science Foundation. Any opinions, findings and conclusions or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

NCAR UCAR | National Center for Atmospheric Research

Outline

1. A tour of the DART website, including how to download DART.
2. A tour of the DART software.
3. How to configure & run DART.
4. Being *FEARLESS* with it – *subversion!*
5. Diagnosing what went right.
6. Diagnosing what went wrong.
7. Common mistakes.
8. Some things to think about.
9. Where to learn more.

www.image.ucar.edu/DAReS/DART

Time for a quick tour of www.image.ucar.edu/DAReS/DART

Things to think about :

1. Registering for DART & getting the code.
2. Where is the documentation?
3. API vs. User Guide
4. How do I use DART?
   – configuring, building, running, testing.

The Organization of DART

Recap:

1. Registering for DART & getting the code.
2. Where is the documentation?
3. API *vs* User Guide

Now …

4. How do I use DART?

cd ~/DART

The DART Graphical User Interface:

`<unixprompt> %`

1. Simple, efficient, clean, script-able …
2. Works on almost any cluster, supercomputer …
3. No need for additional software install/maintenance …

fearless: svn

- Subversion is a version control system that allows you to recover any specific version of a file.
- You can even *delete* the file and get it back.
- If I make an improvement to that file, you can update your file without losing your local modifications!
- You can do a lot even without being 'online'.

Time to destroy DART/models/clm/model_mod.f90

Running an experiment:

Gathering the pieces (this is not an exhaustive list, BTW):

- What parts of the model state need to be part of the DART state?
- Do you have observations? (seems fairly obvious)
  Real or synthetic?
- What kind of cutoff radius do you expect to need?
- Do you have an existing ensemble –or–
  Do you need to get/create one?
- Does your initial ensemble have enough spread? Too much?
- Is your ensemble large enough?
- Does your model need forcing files (ancillary data)?
  Will you need unique forcing files for each ensemble member?
- Is your model MPI-aware or single threaded?
- Does your cluster have a queueing system (PBS, LSF, etc.)?

It ran, but did it work?

Diagnostics … two broad classes:

1. Comparisons with observations – always possible.
2. Comparison of the model state – with something…

I have whole presentations on diagnosing the performance of the assimilation, selecting the cutoff, etc.,

But first, a quick (30 second) recap of how and why we can focus on observation-space diagnostics.

Schematic of an Ensemble Filter for Geophysical Data Assimilation

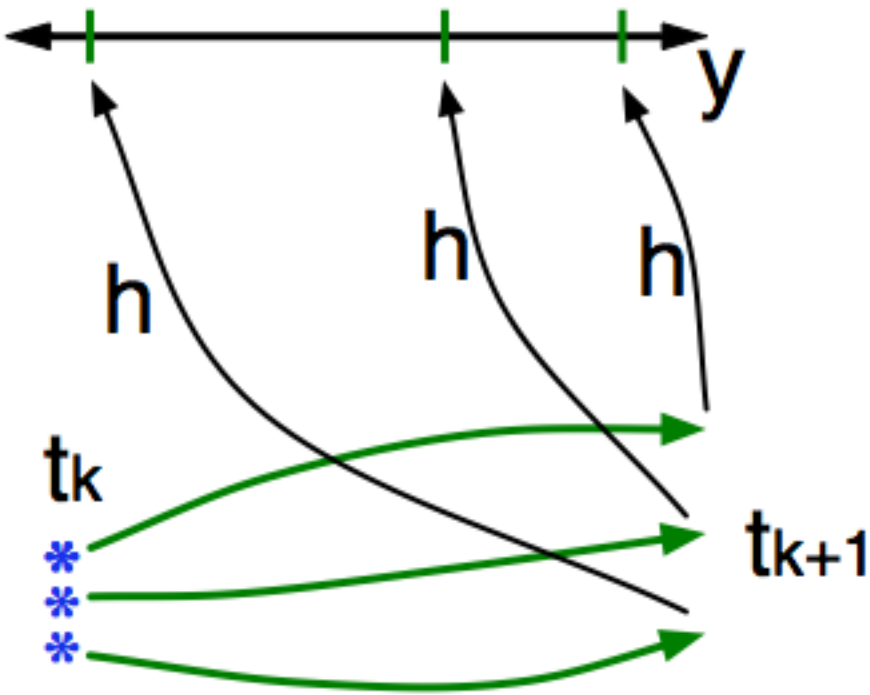
1. Use model to advance ensemble (3 members here) to time at which next observation becomes available.

Ensemble state estimate after using previous observation (analysis)

Ensemble state at time of next observation (prior)

Schematic of an Ensemble Filter for Geophysical Data Assimilation

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

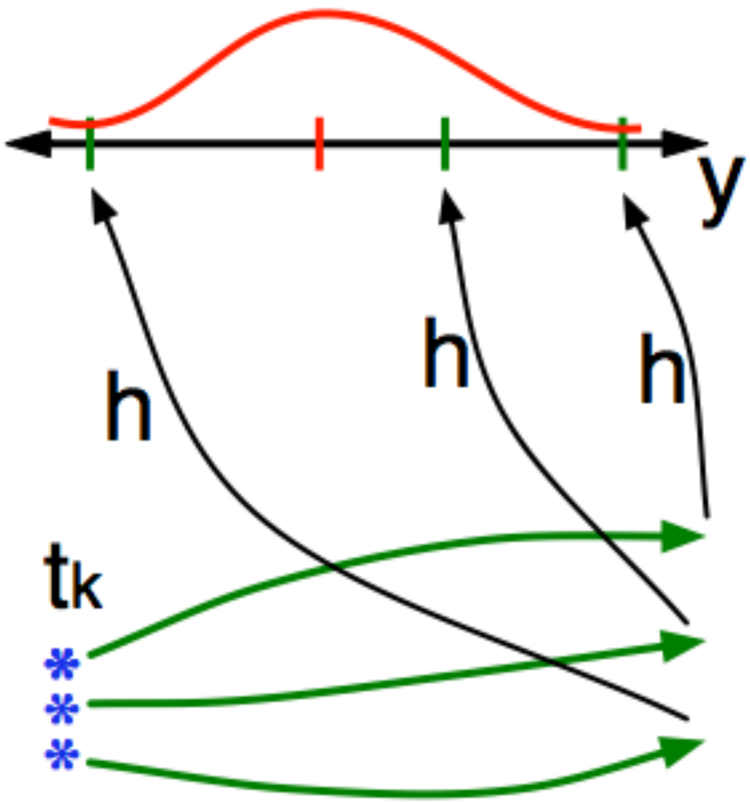


Theory: observations from instruments with uncorrelated errors can be done sequentially.

Houtekamer, P.L. and H.L. Mitchell, 2001: A sequential ensemble Kalman filter for atmospheric data assimilation. *Mon. Wea. Rev.*, **129**, 123-137

Schematic of an Ensemble Filter for Geophysical Data Assimilation

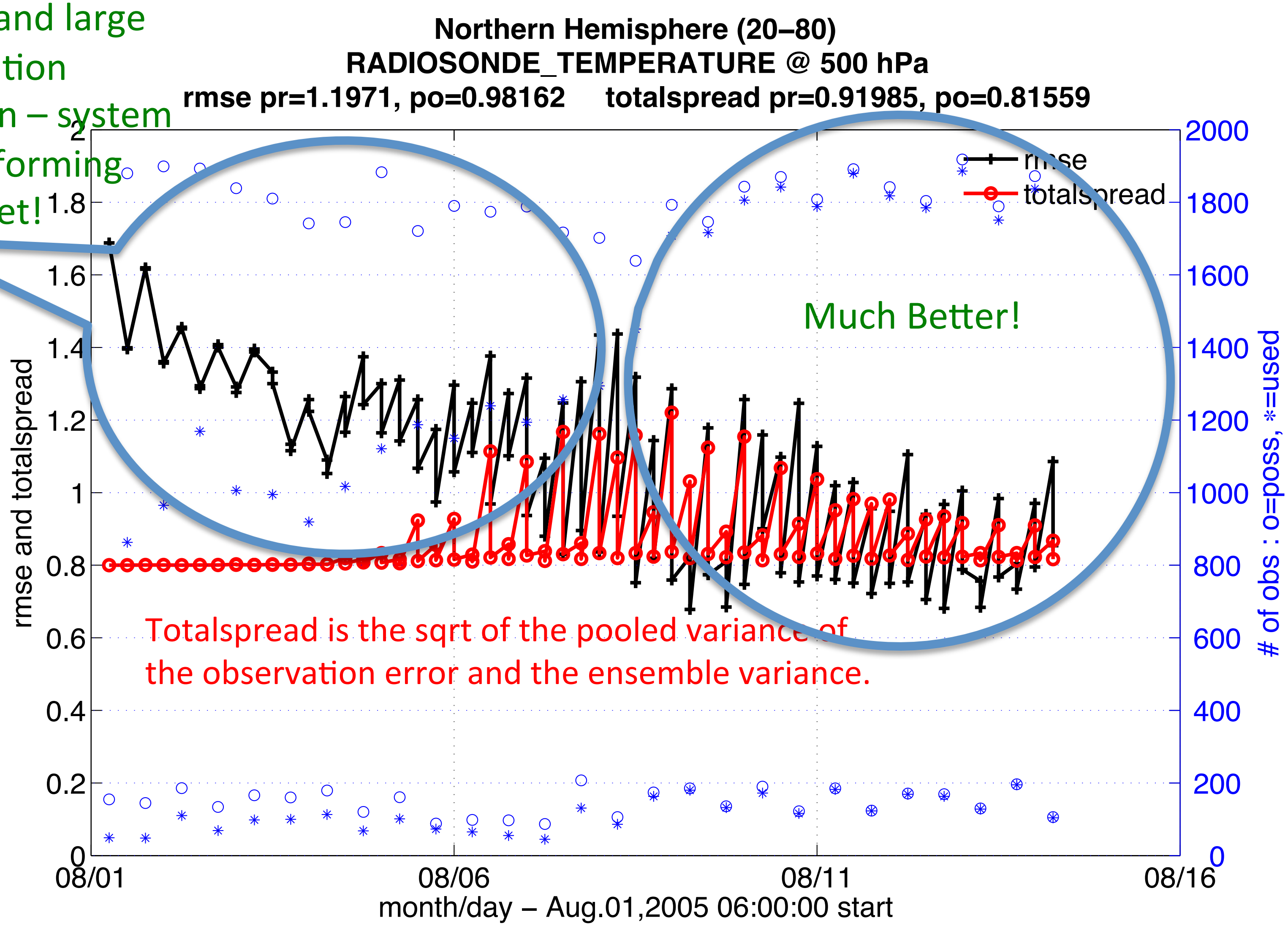
3. Get observed value and observational error distribution from observing system.



If you compare now – i.e. using the priors, you get to compare against 'withheld' observations. You may or may not choose to use them in the assimilation – via the DART namelist. Furthermore, since this is an integral part of the assimilation algorithm, it is computationally **FREE**.

Performance and Rejection

Initially large spread and large observation rejection – system not performing well – yet!

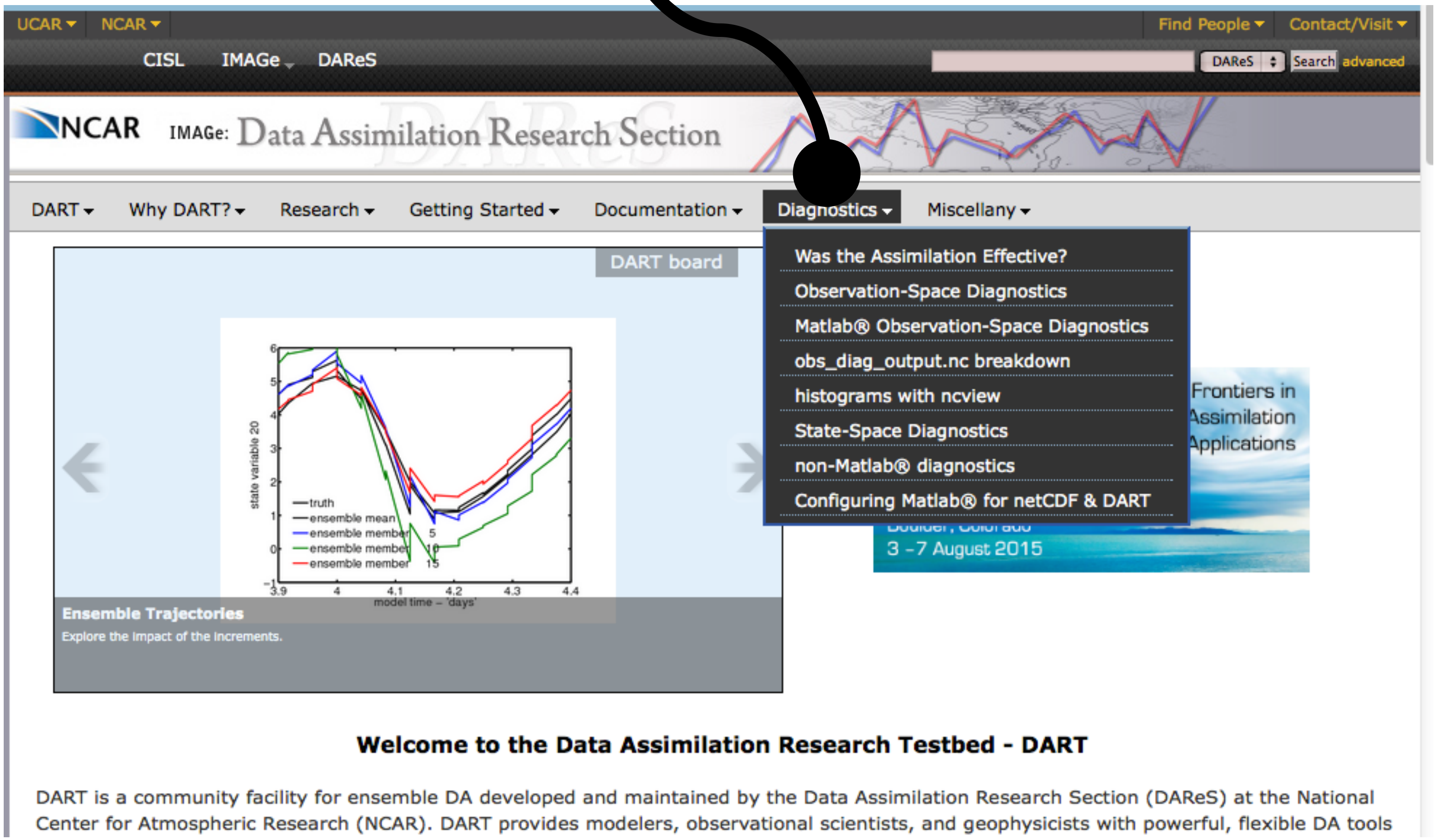


May help explain why analysis is not as good as it could be.

# of observations available AND # actually used!

Where to find DART diagnostic info:

www.image.ucar.edu/DAReS/DART/DART_Documentation.php#DidItWork



DART Matlab functions – in no particular order

State-space: DART/matlab

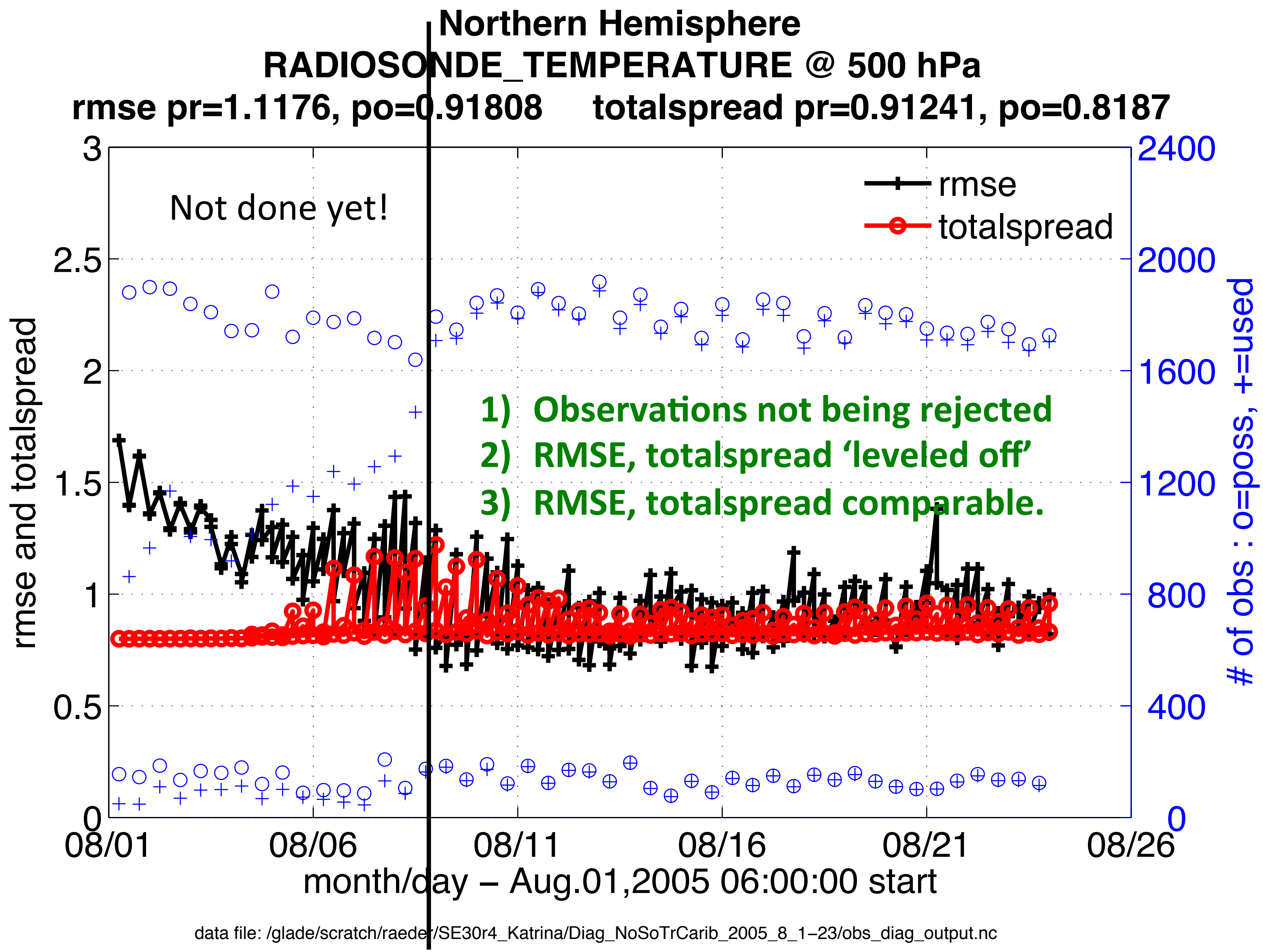
1. plot_bins.m
2. plot_correl.m
3. plot_ens_err_spread.m
4. plot_ens_mean_time_series.m
5. plot_ens_time_series.m
6. plot_phase_space.m
7. plot_reg_factor.m
8. plot_sawtooth.m
9. plot_smoother_err.m
10. plot_total_err.m
11. plot_var_var_correl.m

Observation-space: DART/diagnostics/matlab

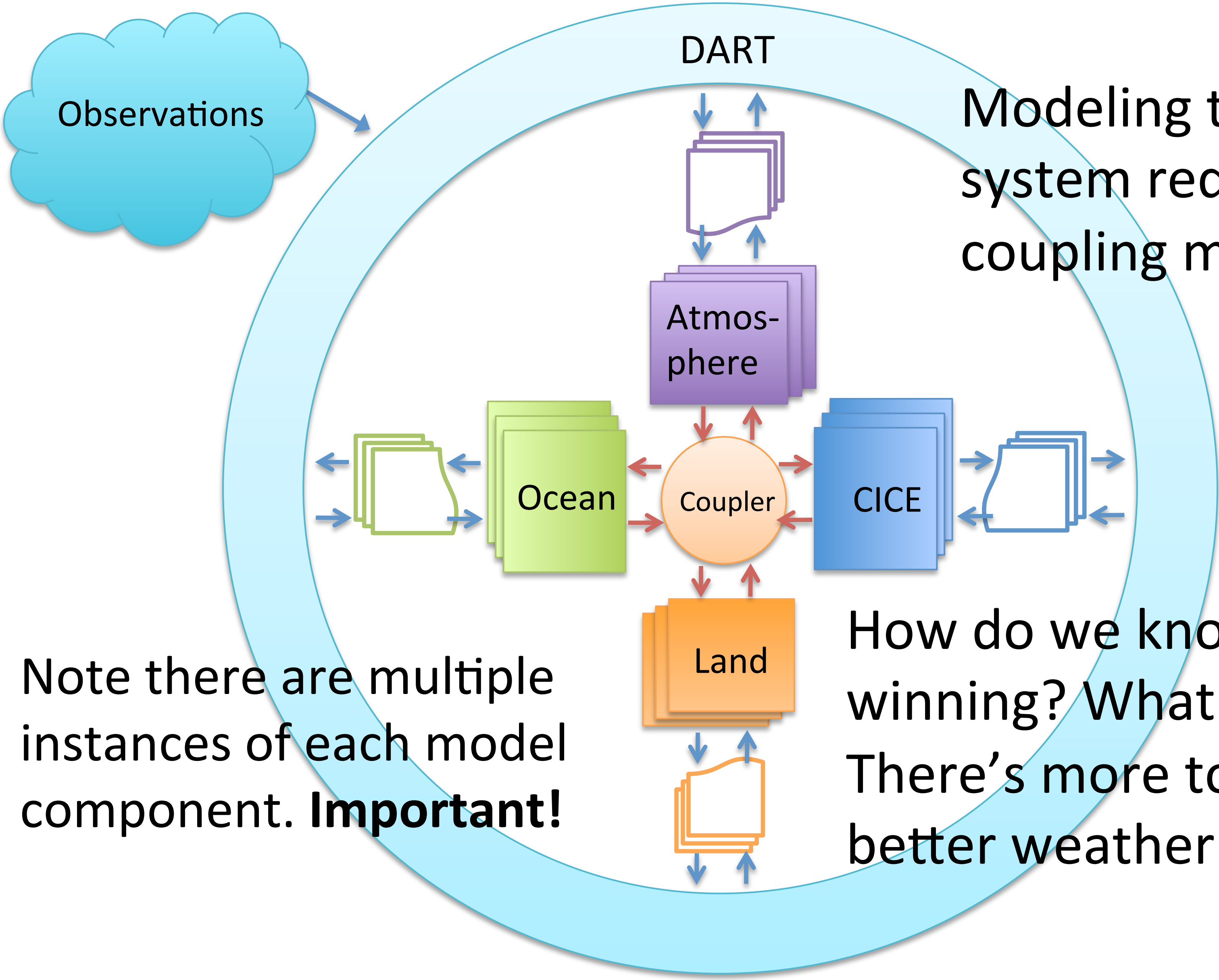
1. plot_bias_xxx_profile.m
2. plot_coverage.m
3. plot_evolution.m
4. plot_obs_netcdf.m
5. plot_obs_netcdf_diffs.m
6. plot_observation_locations.m
7. plot_profile.m
8. plot_rank_histogram.m
9. plot_rmse_xxx_evolution.m
10. plot_rmse_xxx_profile.m
11. plot_wind_vectors.m

If you want to donate your diagnostics – ***in any language*** – we will be happy to give you credit and redistribute them. Since we wind up fielding questions about them, do not be surprised if we bomb-proof them if we can. **Ideas are always welcome!**

A good-looking experiment.



Sometimes the models are *PRETTY COMPLEX*



Modeling the Earth system requires coupling models.

Note there are multiple instances of each model component. **Important!**

How do we know if we are winning? What is success? There's more to it than a better weather forecast!

So … how do we assess performance?

1. We are trying to achieve an ensemble that is indistinguishable from the physical realization of the modeled system. (we want our ensemble of models to generate synthetic observations that have the same PDF as the real observation).

2. We want the ensemble to be as informative as possible (statistical notion of ‘precision’?) and still capture our uncertainty in the system.

3. It is trivial to develop a method to have a terrific **posterior** RMSE compared to observations. ‘Direct replacement’. This was done in the early days of atmospheric DA and it was shown to have **really poor** forecast properties.

4. It is also possible to get a great RMSE by rejecting all the observations that disagree with your ensemble. This is called ‘filter divergence’ and is the #1 undesirable property of ensemble methods. “Show your work.” – how many observations can you use, how many did you use?

Rank histograms can assess #1 and #2.
Observation-space diagnostics of the ***PRIOR*** can assess #3 and #4.

Minimal list of things to assess:

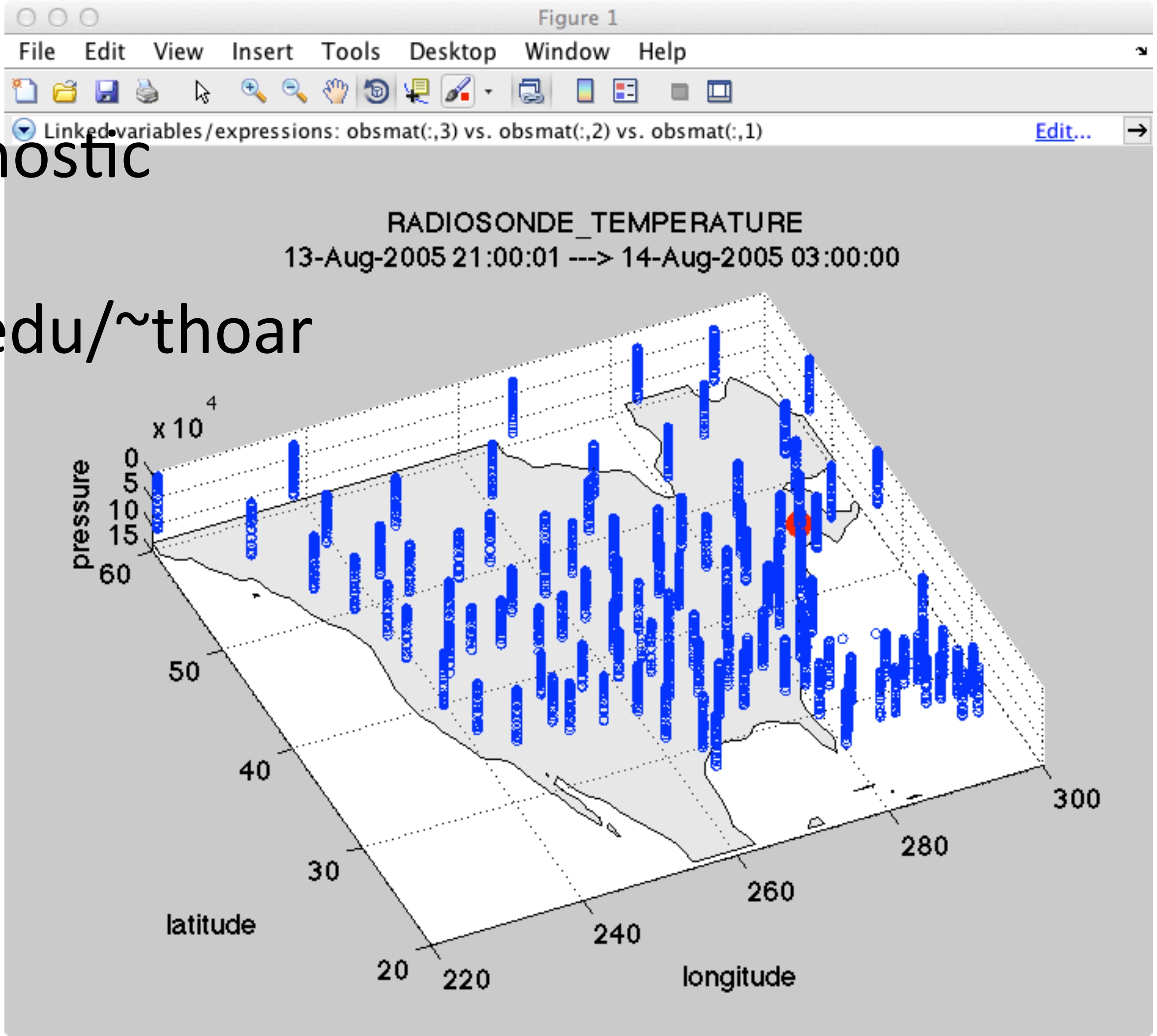
In no particular order:

1. Are the observations getting rejected?
2. Is the ensemble collapsing?
3. Is the RMSE more-or-less steady?
4. Do the rank histograms look reasonable?
5. If you are using inflation, are the inflation values reasonable?
6. Is the model state reasonable?

There are many more application-specific metrics … depends on your objective.

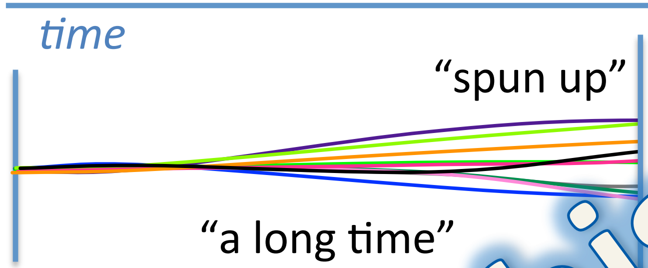
Rejection … where and why?

Get the CAM diagnostic files from www.image.ucar.edu/~thoar



Tim – don’t forget to run Matlab GUI **link_obs** for `models/cam/work`

More things to think about:



Getting a proper initial ensemble is an area of active research.

Does this work for your model?

1. Replicate an equilibrated state N times.
2. Use a unique (and different) *realistic* forcing for each to induce separate model trajectories.
3. Run them forward for “a long time”.

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

The HARD part:

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"?
Crusty Snow?
Slushy Snow?

Old Snow?
Dirty Snow?
Packed Snow?
Early Season 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. The land and chemistry models have hundreds of carbon-based quantities!

Time to run DART: lorenz_96

`<unixprompt>: cd models/lorenz_96/work`

1. Get familiar with the files …
   - *mkmf_* and *path_names_* … pairs?
   - data files
   - observations
   - Scripts
2. Open them, look at them, READ THE COMMENTS, modify them, use subversion to compare, use subversion to revert them … i.e. be inquisitive and fearless.

Examine what happens in *workshop_setup.csh*

Time to run DART: lorenz_96

```
<unixprompt>: cd models/lorenz_96/work
```

1. Did the experiment work …
   - Run *plot_total_err.m*
   - Run ./***obs_diag*** and then *plot_evolution.m*
   - Examine the rank histograms … I am not telling you how – we have covered it.

2. What do you think you could do next? Not a rhetorical question – this is the part where you get to explore!

Tim – set a timer to leave 5 minutes till the end.

Key Questions for Ensemble DA:

- What parts of the model ‘state’ do we update?
- What is a proper initial ensemble?
- Is an ensemble of boundary conditions necessary?
- Localization considerations
- How many ensemble members are needed? to mitigate regression error?
- What is the proper observation error specification? It is not just instrument error but also mismatch in representativeness.
- Can models tolerate new assimilated states? Silently fail? Violently fail?
- Forward observation operators
  - Many observations are over timescales or are quantities that are inconvenient
- Bounded quantities? When all ensembles have identical values the observations cannot have any effect with the current algorithms.

Climate Modeler's Commandments

by John Kutzbach (Univ. of Wisconsin).

1. Thou shalt not worship the climate model.
2. Thou shalt not worship the climate model, but thou shalt honor the climate modeler, that it might be well with thee.
3. Thou shalt use the model that is most appropriate for t ion at hand.
4. Thou shalt not change more than one thing at st.
5. In making sensitivity experiments, thou model hard enough to make it notice you.
6. Thou shalt not covet fine-scale res arse-scale model.
7. Thou shalt follow the rules f ce testing and remember the model's inherent variability.
8. Thou shalt know the model and remember that model biases may lead to biased sensitivity estimates.
9. Thou shalt run the same experiment with different models and compare the results.

Amen Brother!

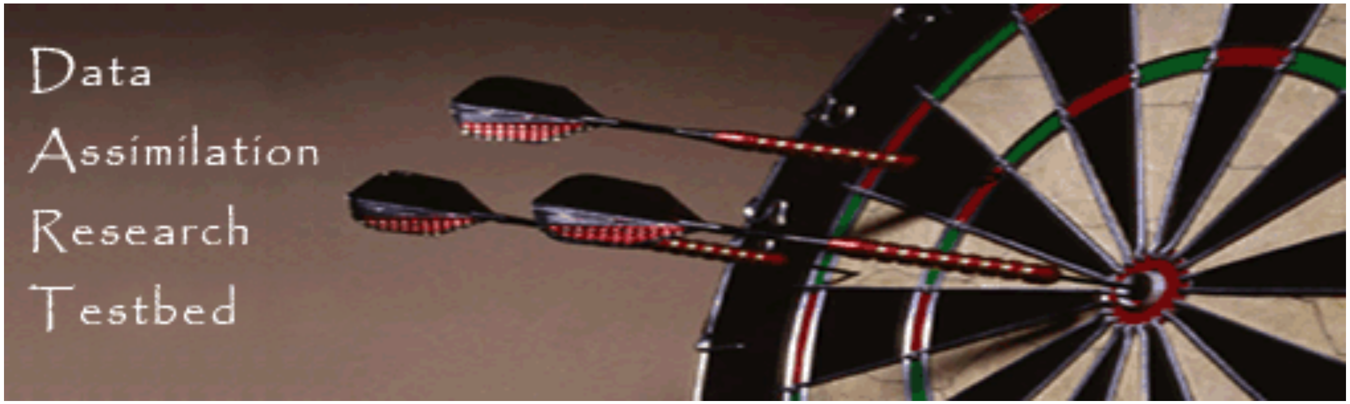
10. Thou shalt worship good observations of the spatial and temporal behavior of the earth system. Good models follow such observations. One golden observation is worth a thousand simulations.

For more information:

CAM GCOM CAM-Chem ROMS WRF

GITM WRF-Hydro WACCM

CLM



POP

AM2

BGRID

COAMPS www.image.ucar.edu/DAReS/DART NOAH

dart@ucar.edu

MPAS_ATM

MITgcm_ocean

SQG NAAPS MPAS_OCN TIEGCM COAMPS_nest

PE2LYR

NCOMMAS PBL_1d

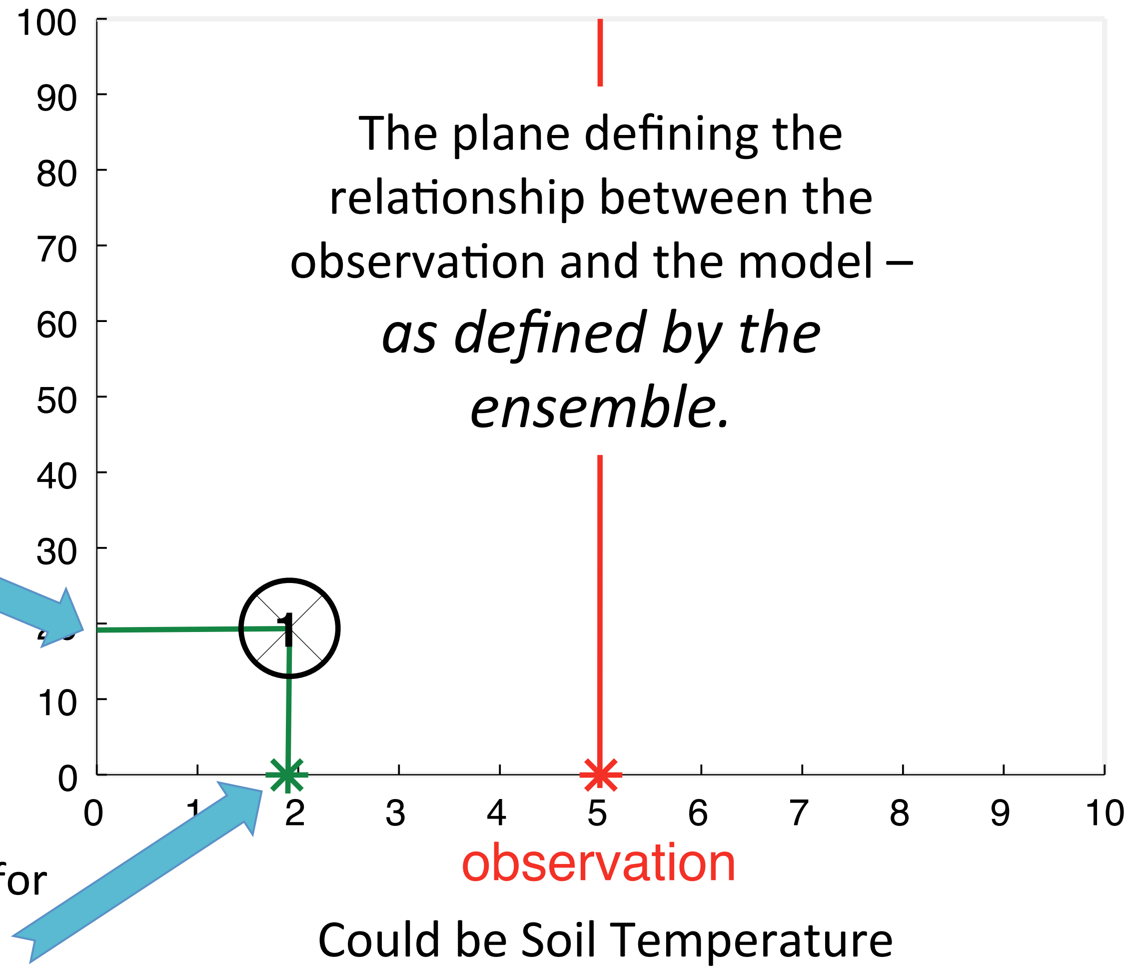
CABLE WRF-Chem

Everything after here held in reserve.

Looking at it another way:

Some unobserved state variable. e.g. live root carbon, dead root carbon, canopy water …

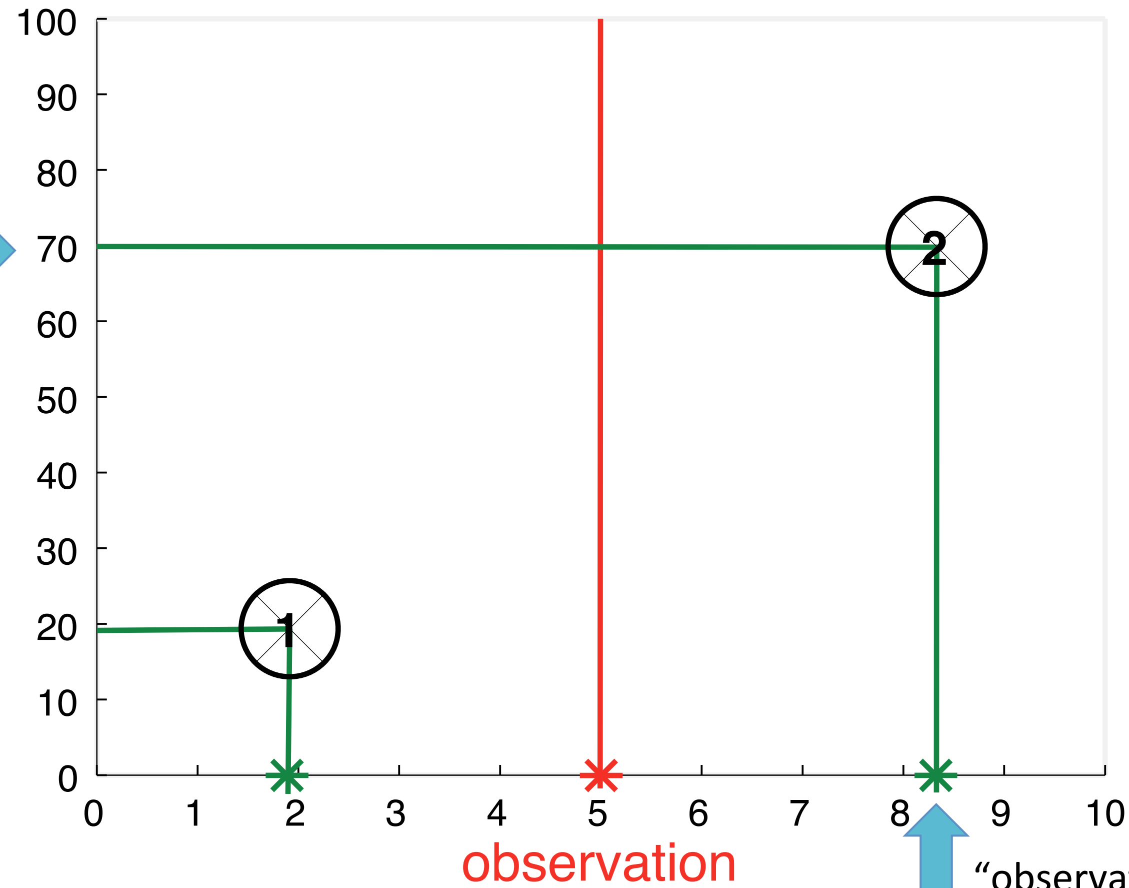
Directly from ensemble member 1



Result of the forward observation operator for ensemble member 1

Looking at it another way:

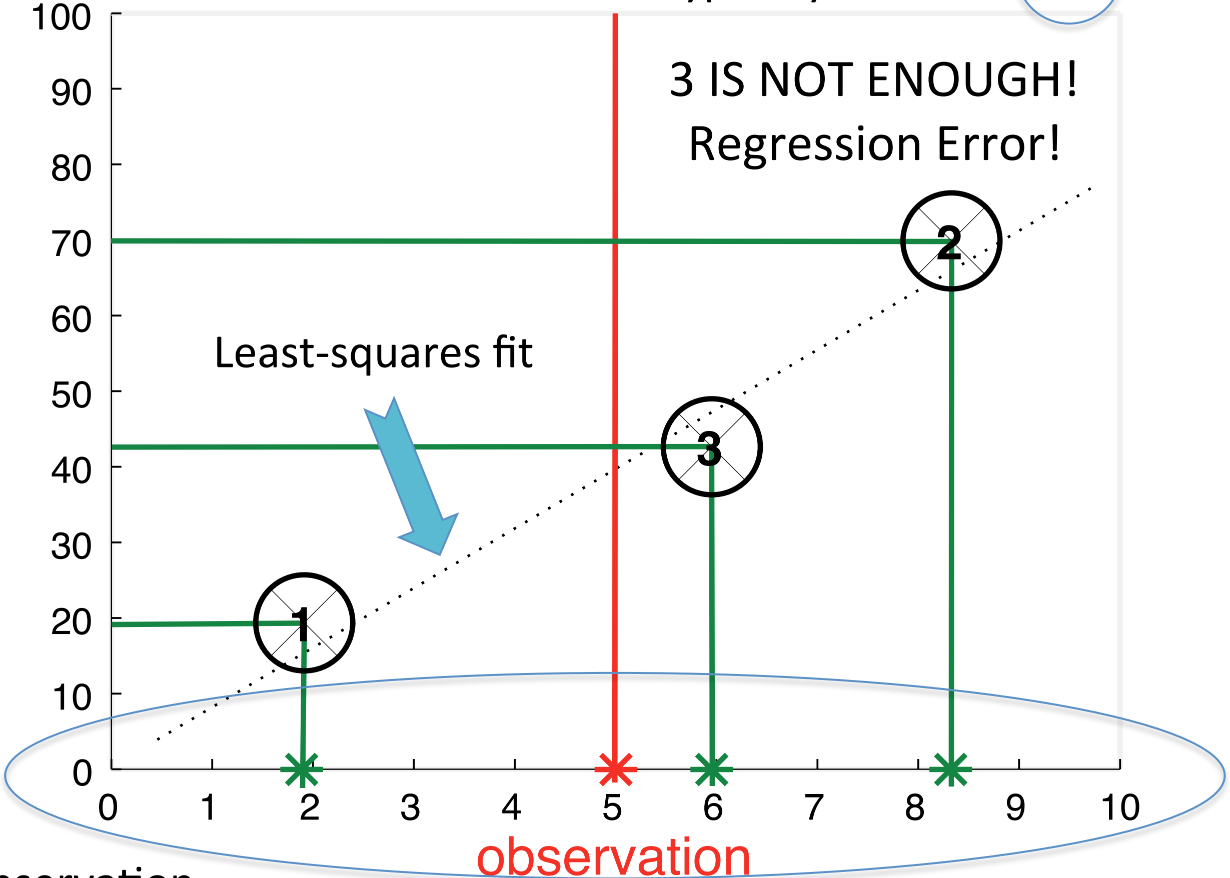
Directly from
ensemble member 2



"observation"
from ensemble
member 2

Looking at it another way:

In our assimilations, we typically use order 80.

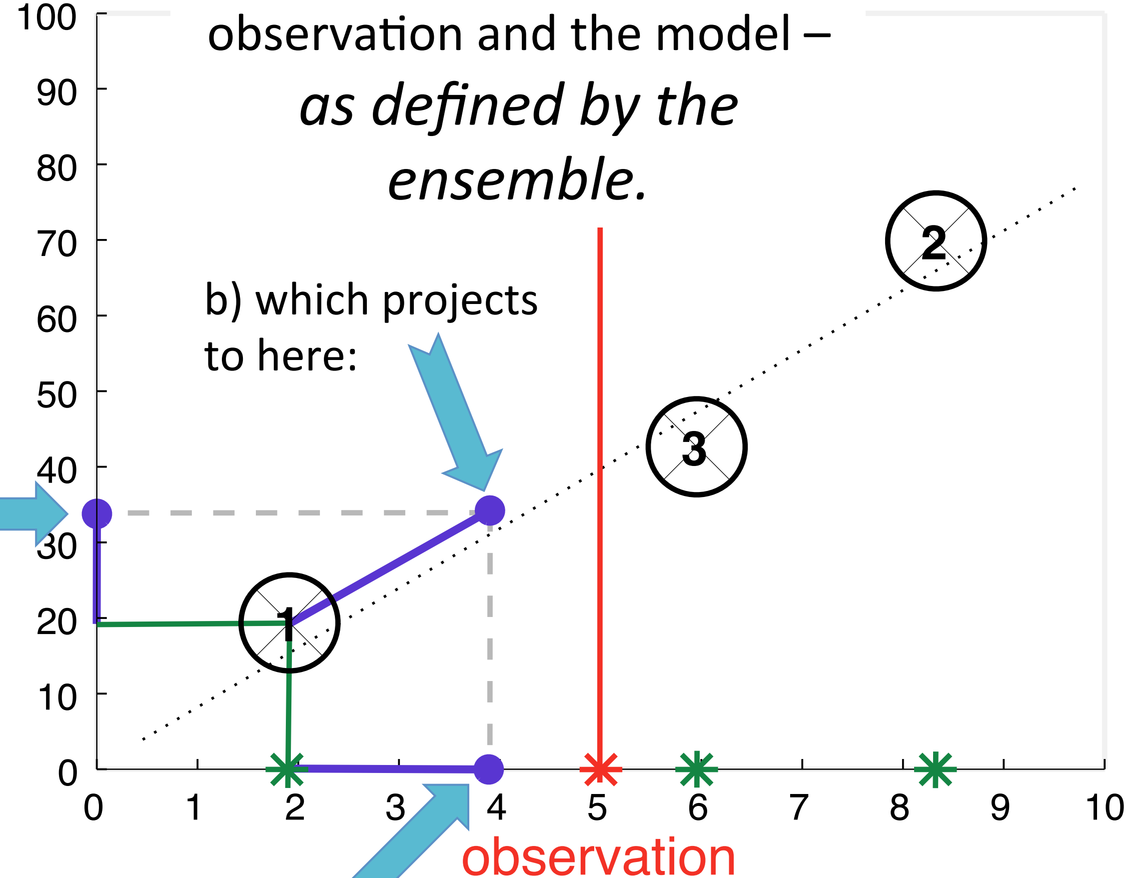


Now, we can calculate out observation increments any way we want.

Looking at it another way:

Anderson, J.L., 2003:
A local least squares framework for ensemble filtering. *Mon. Wea. Rev.*, **131**, 634-642

The plane defining the relationship between the observation and the model – *as defined by the ensemble.*



b) which projects to here:

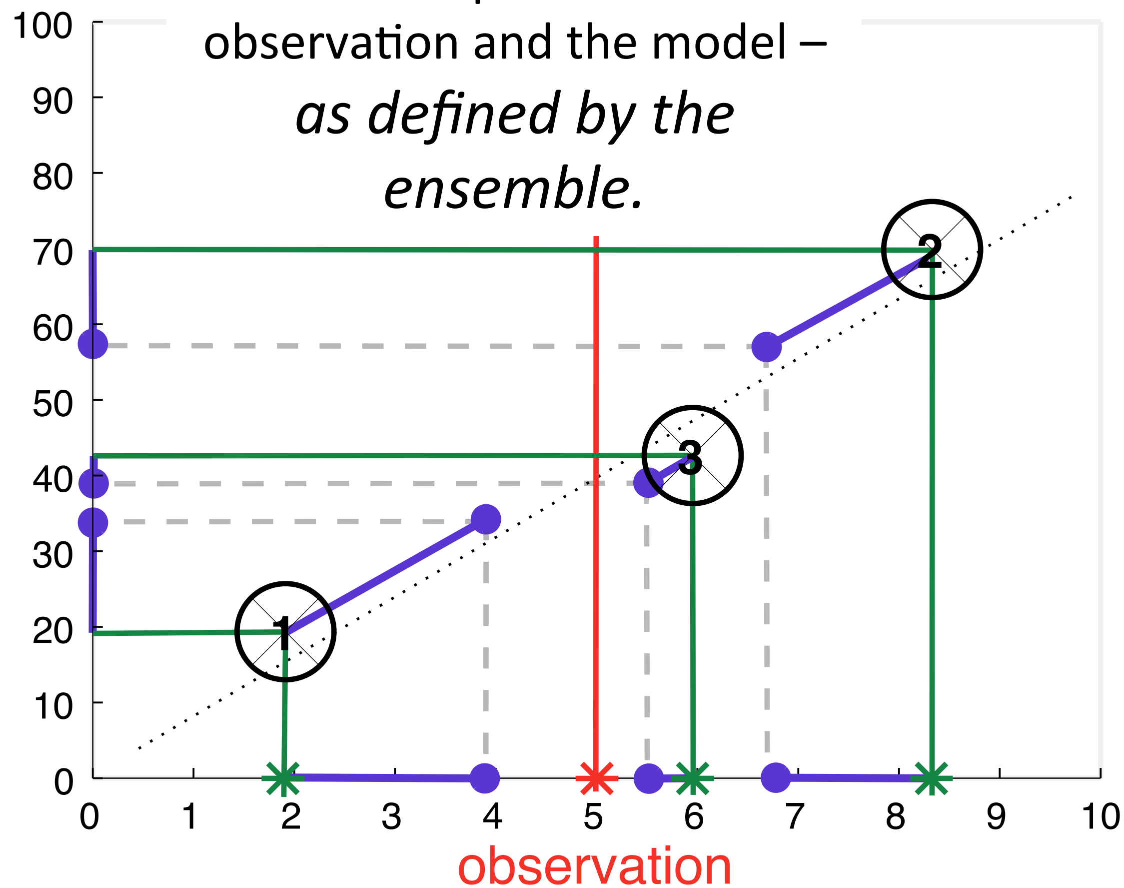
c) Which means the unobserved Posterior should be:

a) The “observation” Posterior for member 1

Looking at it another way:

Any part of the model: snow cover fraction, root carbon, canopy water …
Could even be a model parameter!

The plane defining the relationship between the observation and the model – *as defined by the ensemble.*

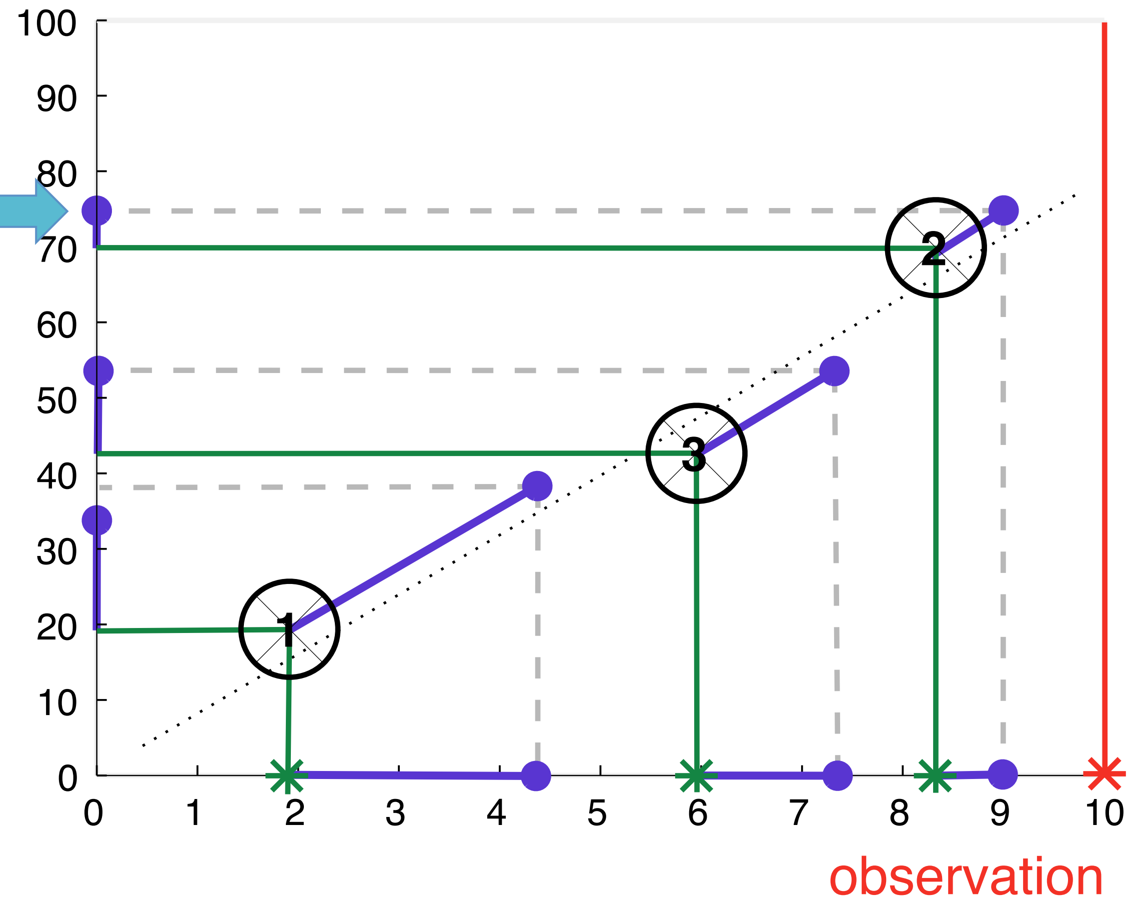


Could be Soil Temperature

Potential Problem

This posterior MAY or MAY NOT be realistic!

Can the model tolerate this new state?



If the observation is "too far" away, it is rejected.

What is "too far"?

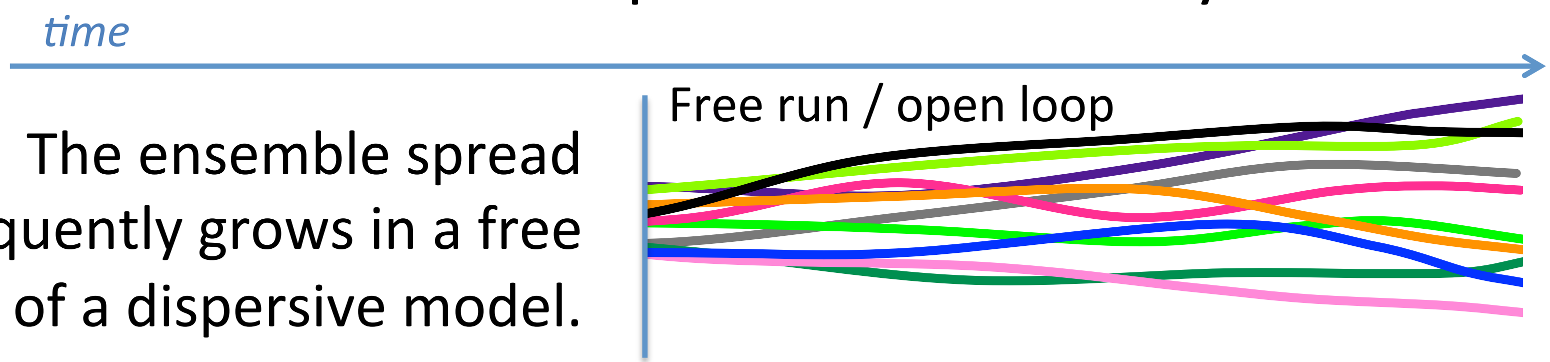
Future Work: AKA “What I didn’t talk about.”

- ✓ Improved observation metadata / peculiar land model hierarchies …
- ✓ Snow … destroying is easy, making ‘brand new’ snow is hard …
- ✓ Forcing files/data for the resolutions desired …
- ✓ Forward observation operators in support of the instruments …
- ✓ Supporting non-local localizations (eg. watersheds) …
- ✓ The initial ensemble & spread …
- ✓ Identifying model variables that *NEED* to be updated …

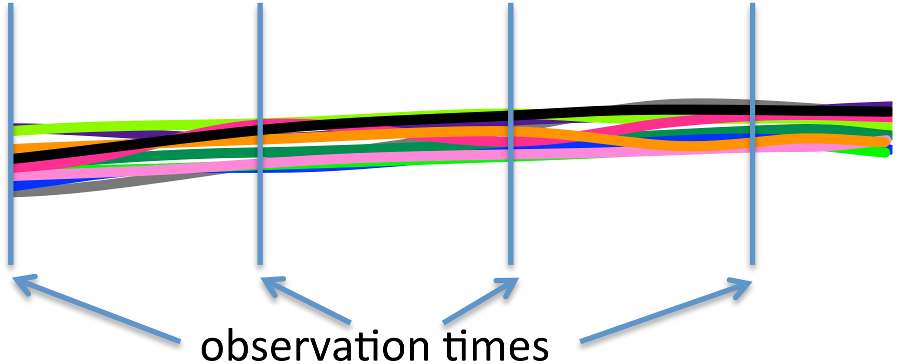
And a whole lot more …

The ensemble advantage.

You can represent uncertainty.

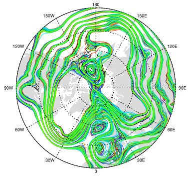


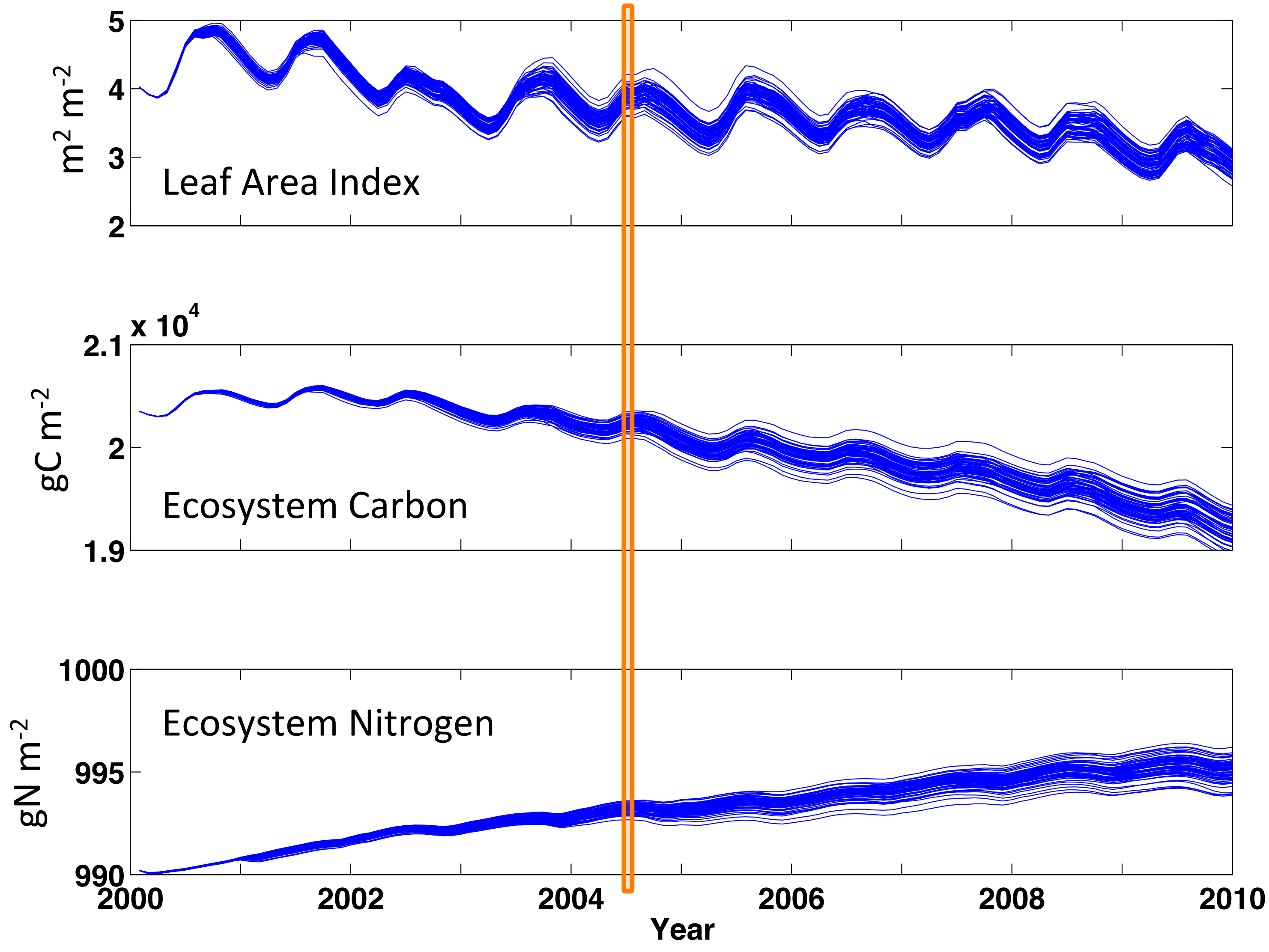
The ensemble spread frequently grows in a free run of a dispersive model.



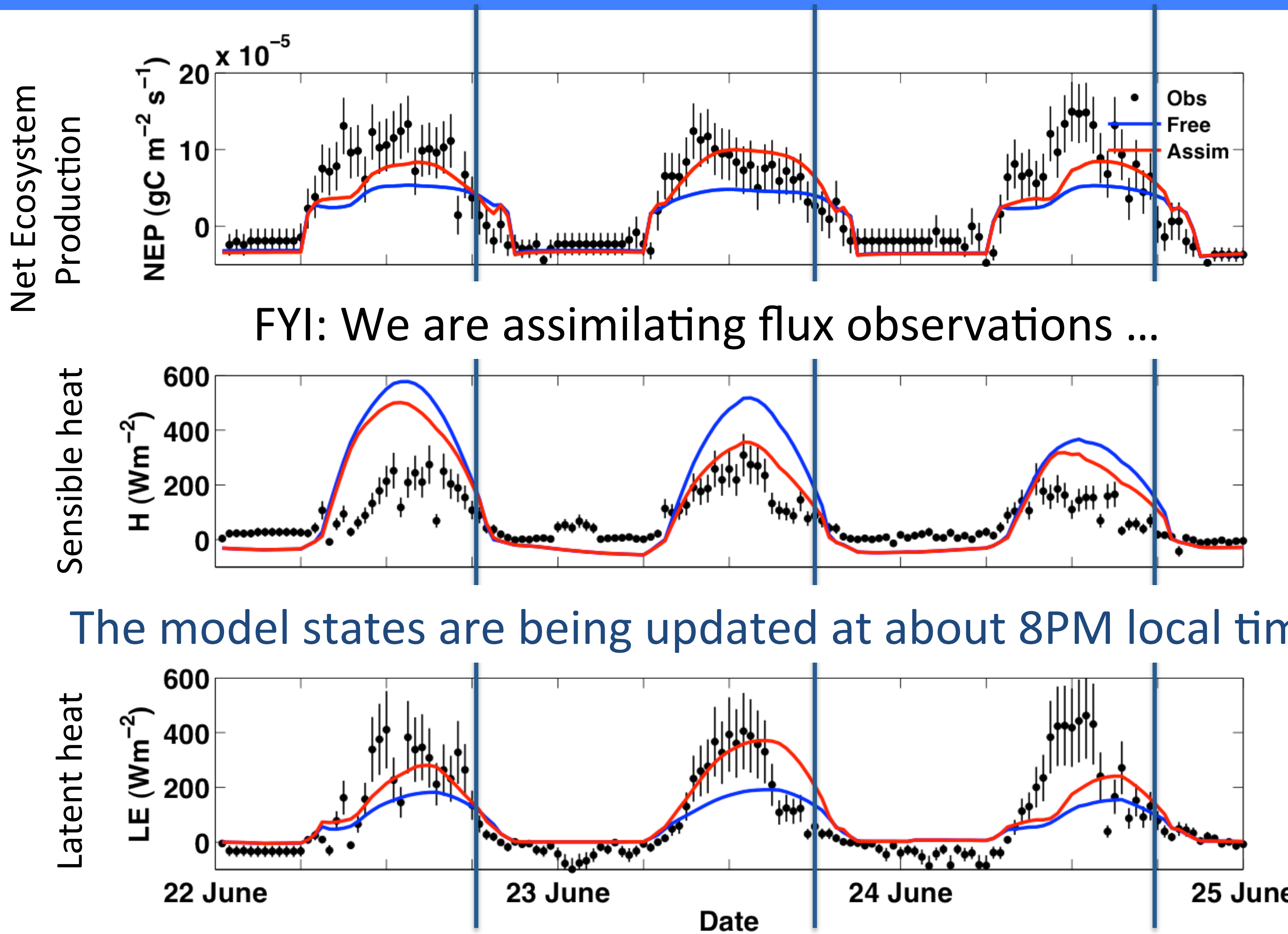
A good assimilation reduces the ensemble spread and is still representative and informative.

Free Runs of CLM driven by 64 CAM reanalyses

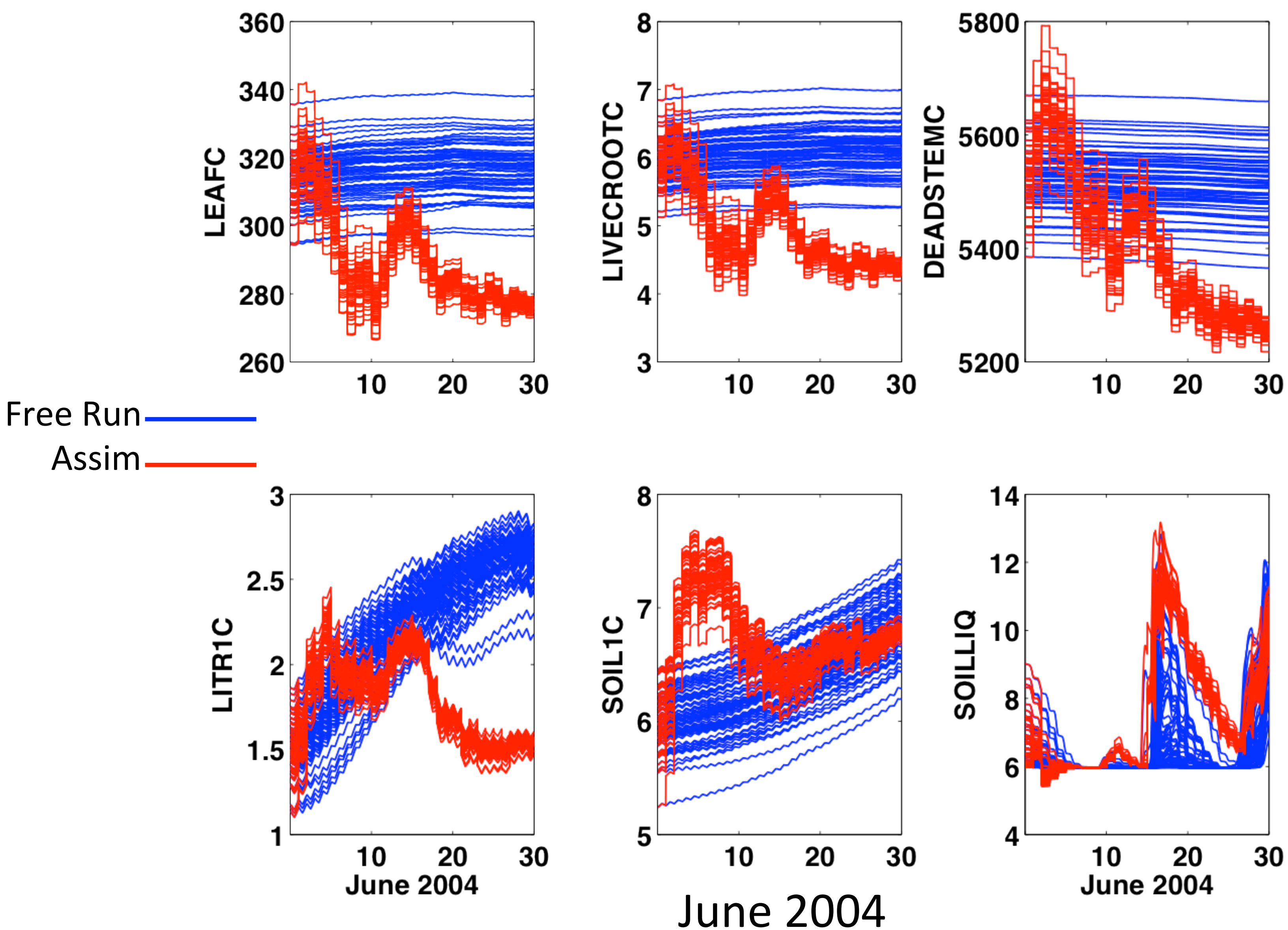




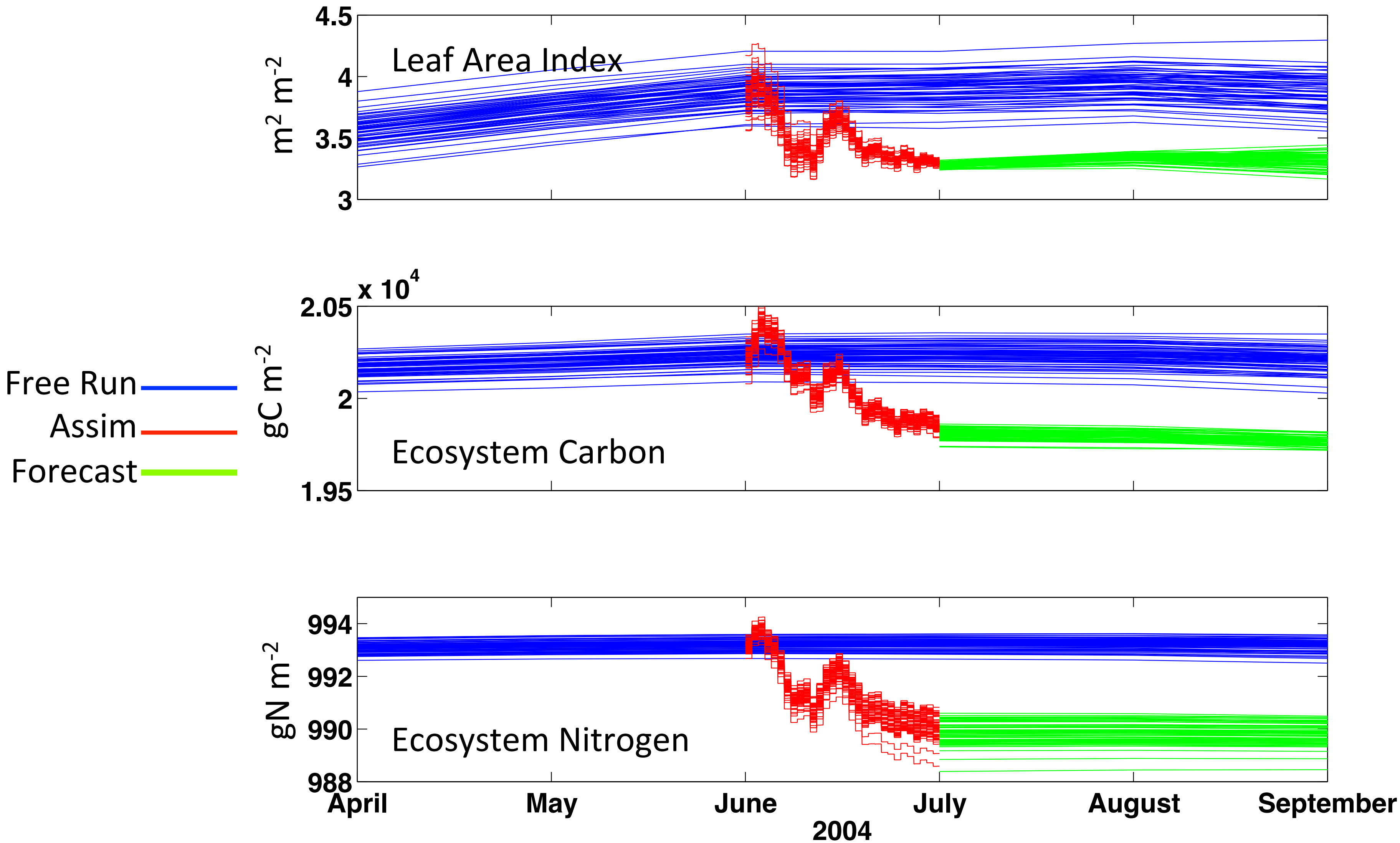
In collaboration with Andy Fox (NEON): Focus on the ensemble means (for clarity)



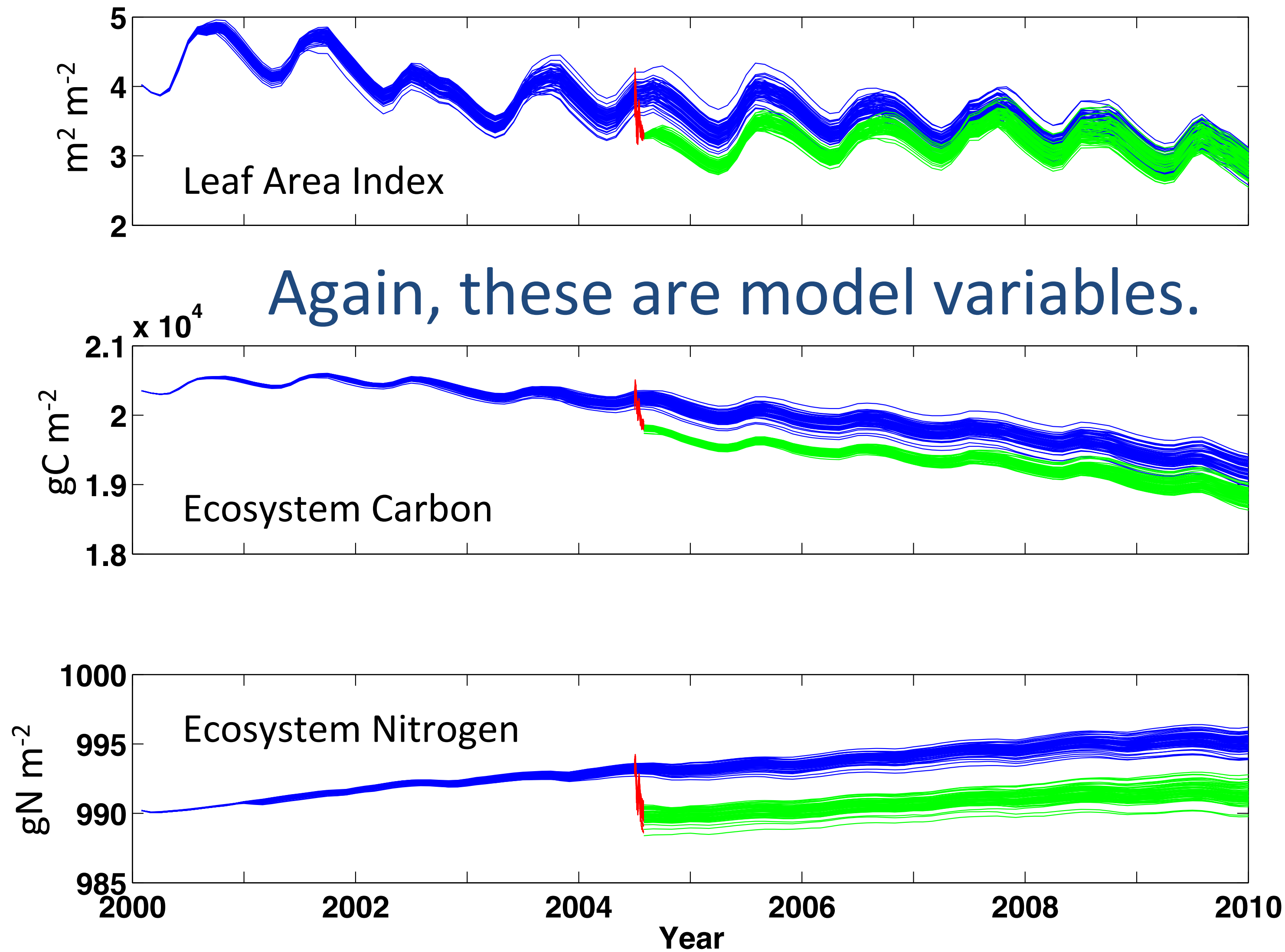
These are all unobserved variables.



Effect on short-term forecast on unobserved variables.



Effect on longer-term forecast



Again, these are model variables.

At the very least : don't compare this:

Your fully-tested, optimized final product.

To this:

Something full of unrealized potential.

Or even more disheartening:

Don’t compare this to this.

It is possible to sabotage (even unintentionally) a method to produce poor results.

Sadly, it happens!