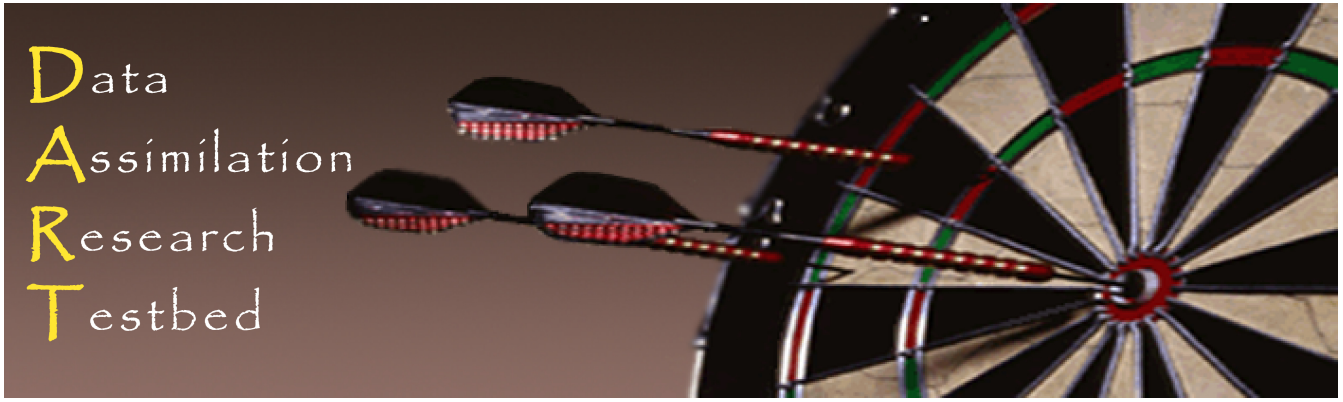
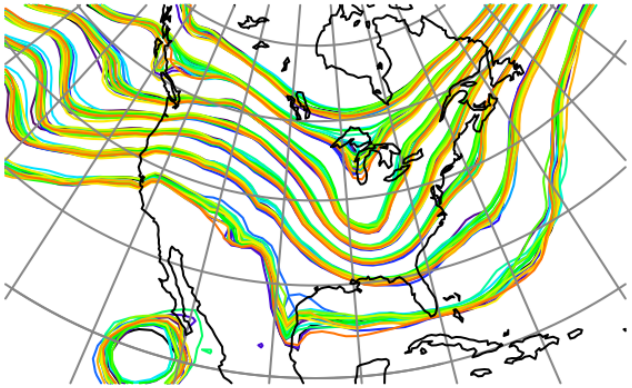


Ensemble Verification: Part II

Tim Hoar, 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.

Ensemble Verification: definition?

The atmospheric science literature is rich with ways to quantify the accuracy of a forecast. I am not going to discuss any of these. Whole conferences could be organized around this topic.

All too often the poster/presentation states "I pushed the EnKF button and here are the results." with little to no regard as to whether or not the assimilation effectively used the information in the observations or whether the ensemble was useful.

I am going to focus on **metrics that inform about the effectiveness of an ensemble forecast assimilation system**.

The Big Questions:

1. Are you looking at the prior (Good!) or the posterior (nearly useless for assimilation assessment – think 'direct replacement')?

2. Do you have an appropriate ensemble?
   - Underdispersed …. filter divergence
   - Overdispersed … uninformative
   - Biased …

3. Are your metrics skewed by observation rejection?

4. Where and why are observations being rejected?

Introduction to DART

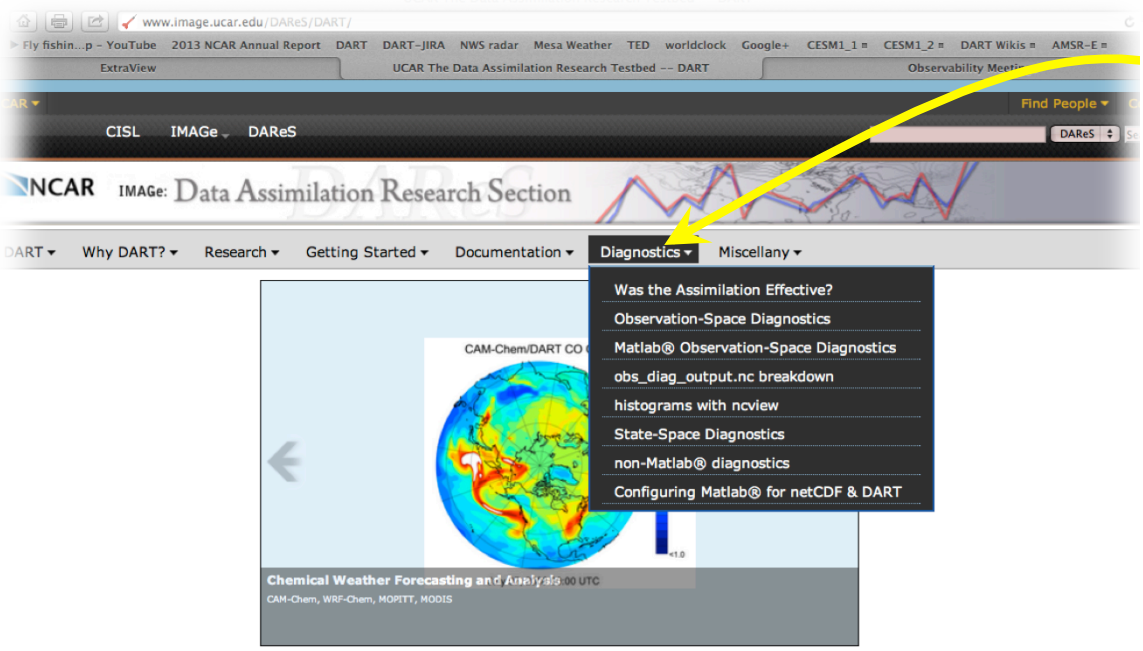
DART is the Data Assimilation Research Testbed, an open source, freely available, ensemble data assimilation system that works with many geophysical and dynamical models and many types of observations.

DART is Fortran-based, scales well into the thousands of processors, has tutorials and documentation, and a small support staff.

We also have tools to evaluate the performance of an assimilation system– as opposed to a forecast – I'm going to use those tools to demonstrate what we believe to be important considerations.

DART "home page":

http://www.image.ucar.edu/DAReS/DART



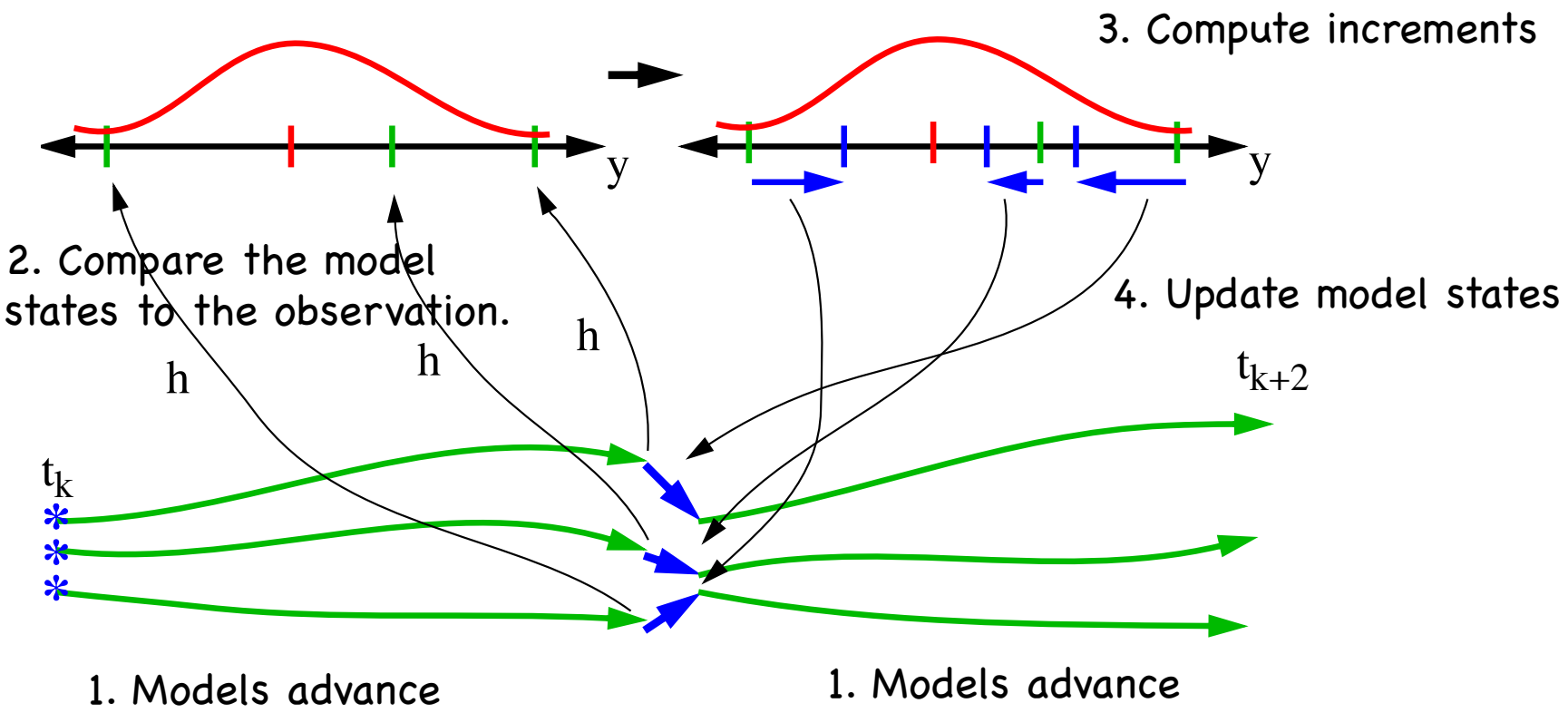
Diagnostics

- Was Assimilation Effective?
- Observation-Space Diagnostics
- …
- Histograms
- State-Space Diagnostics
- …

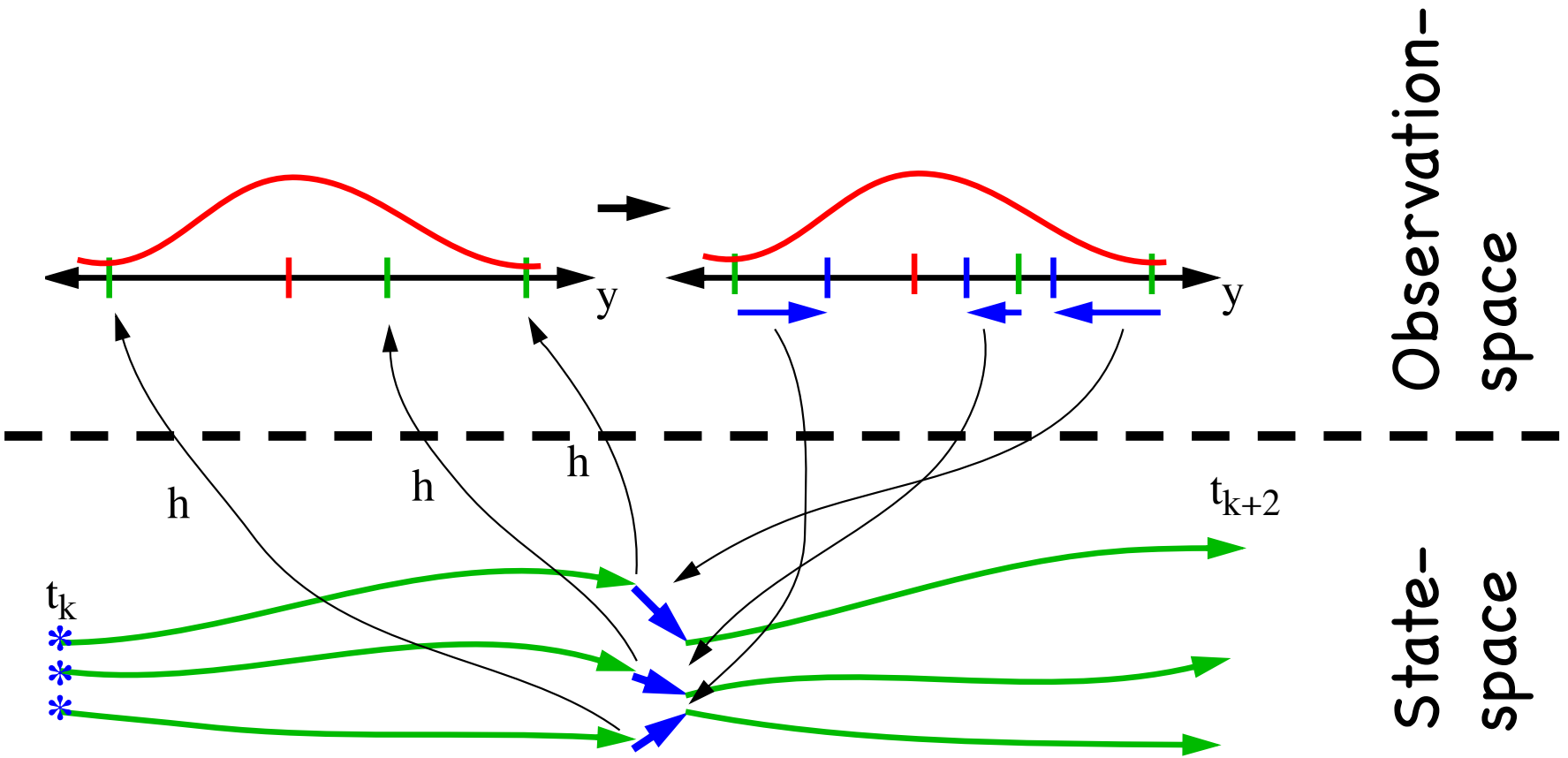
Overview article of DART:

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

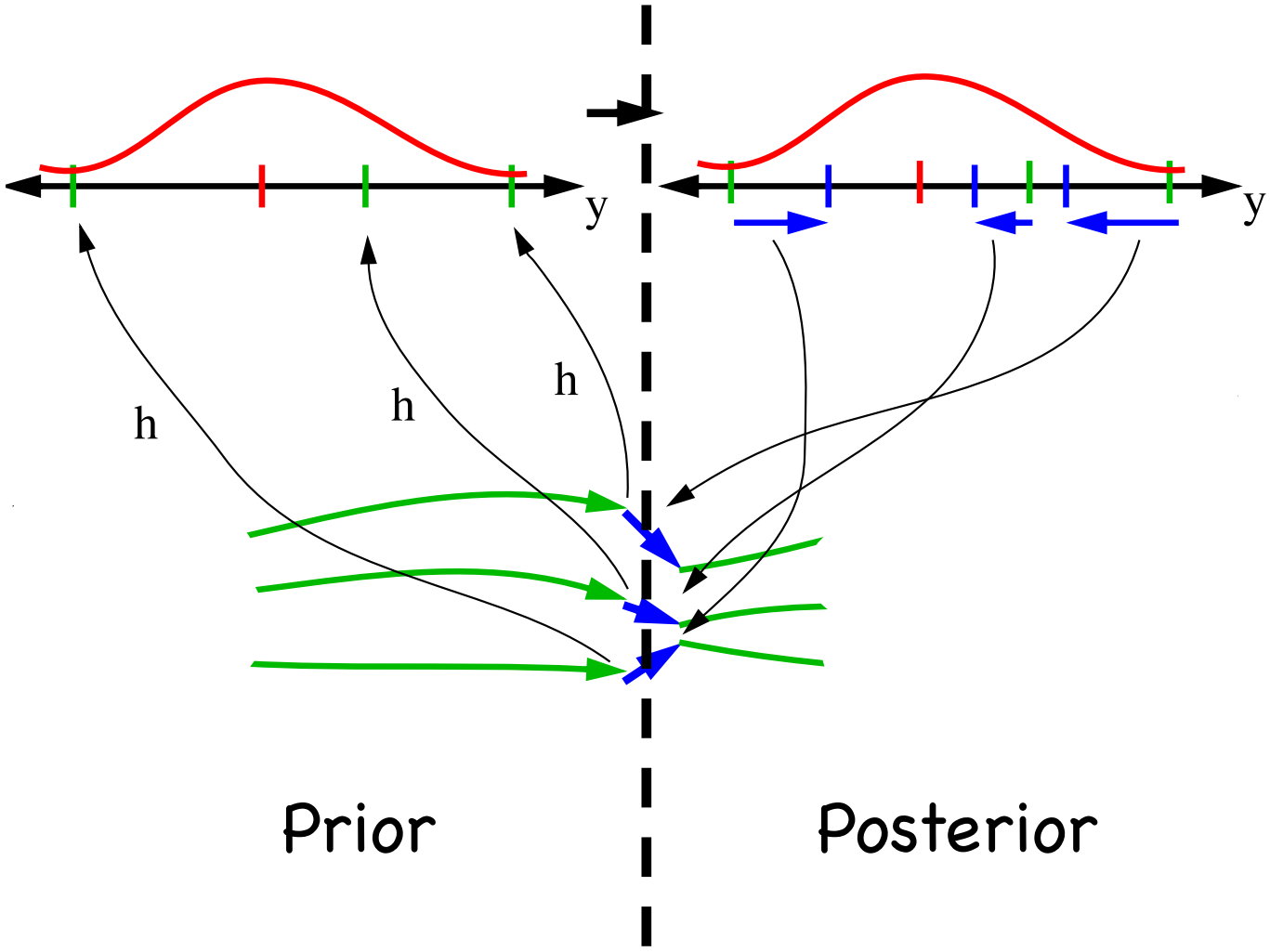
Schematic for any ensemble system:



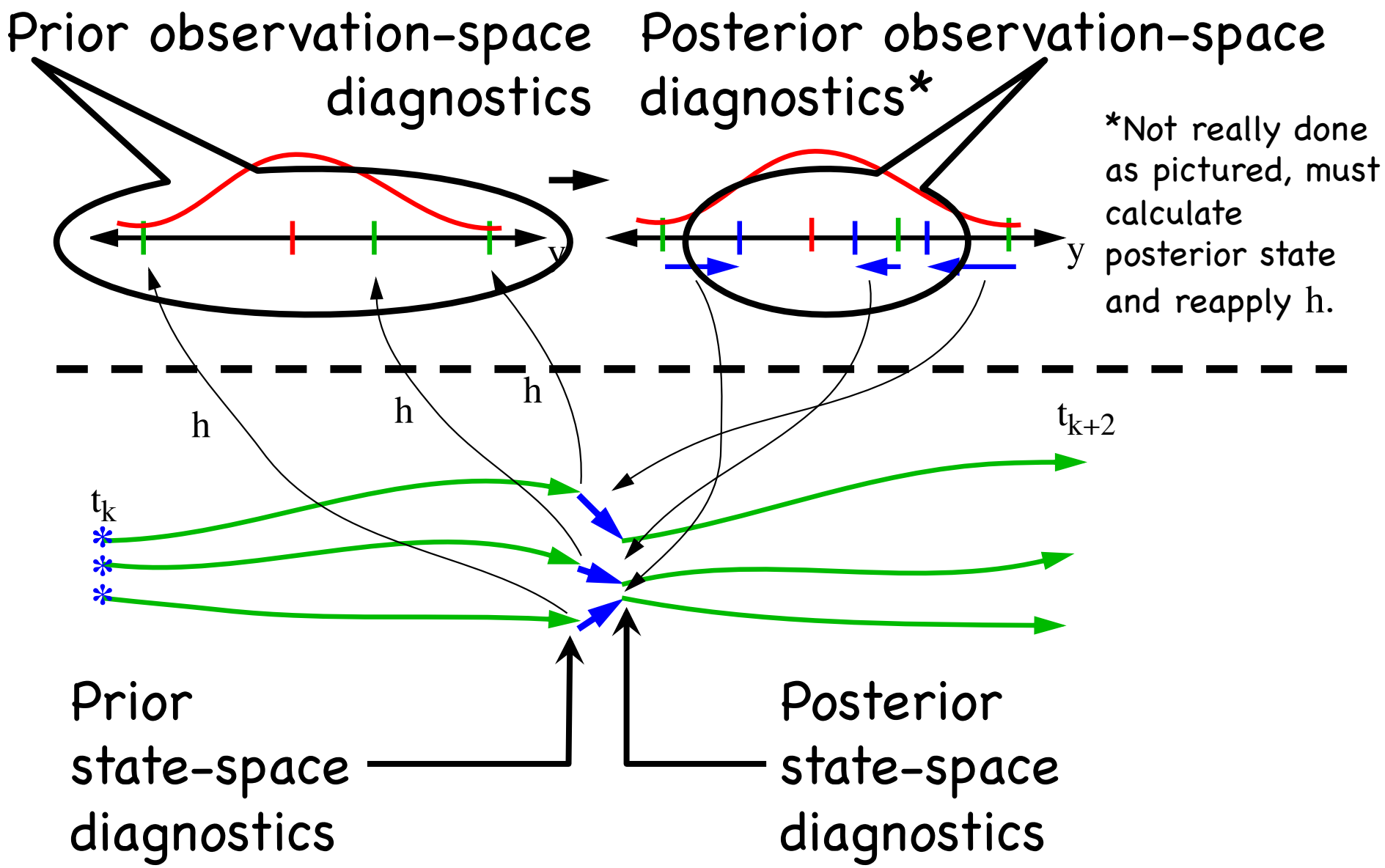
Intro to DART-centric view and tools



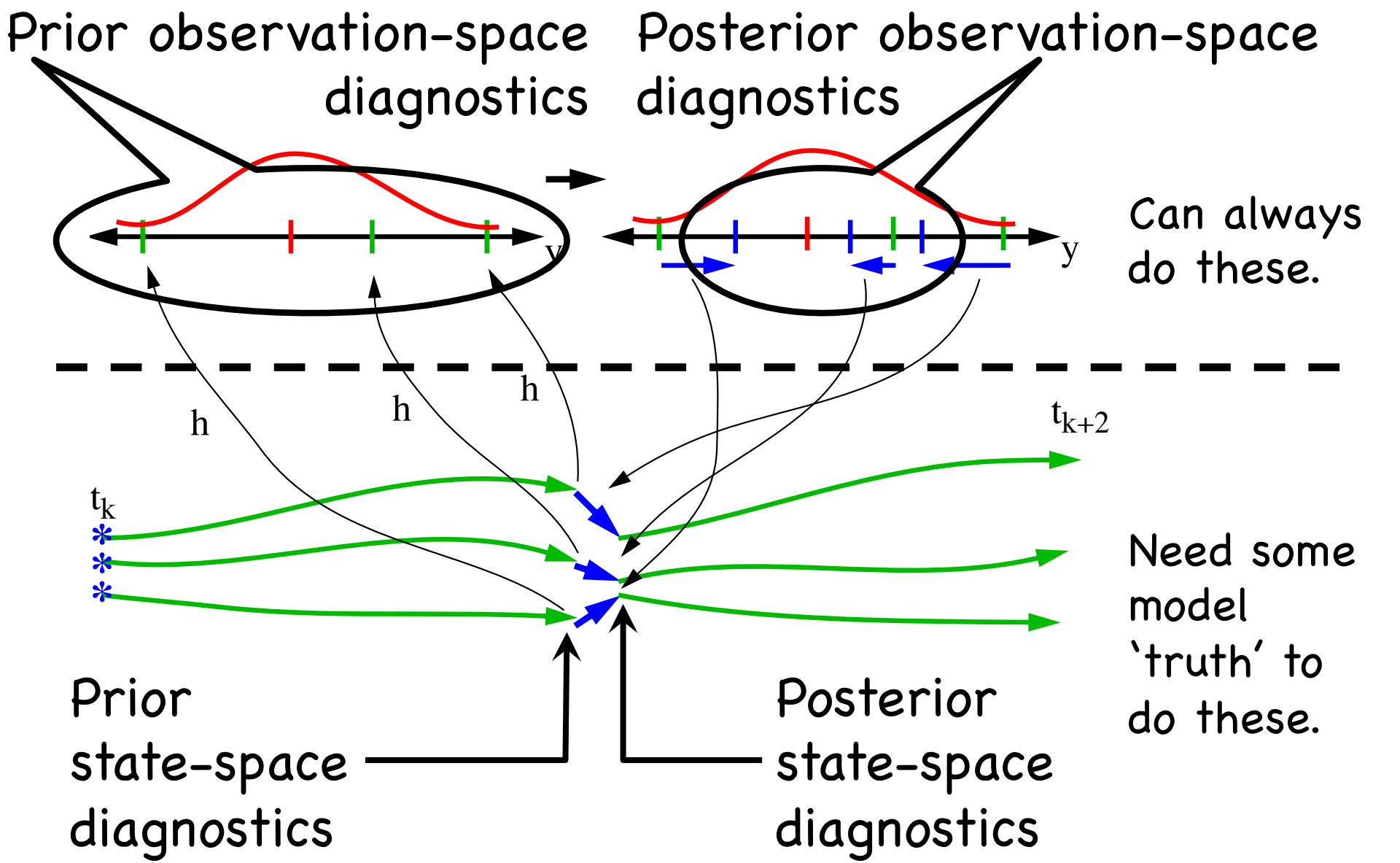
Intro to DART-centric view and tools



Intro to DART-centric view and tools

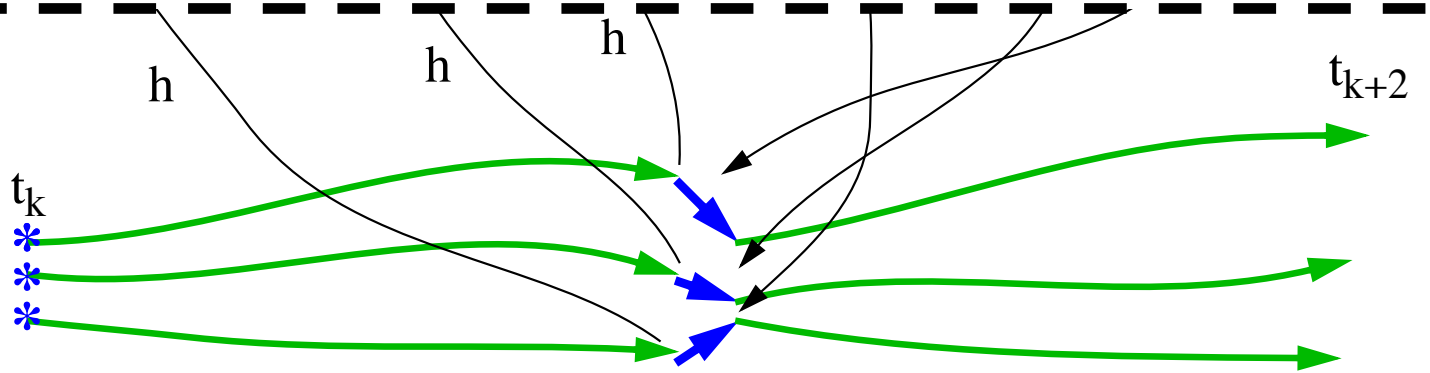


Intro to DART-centric view and tools



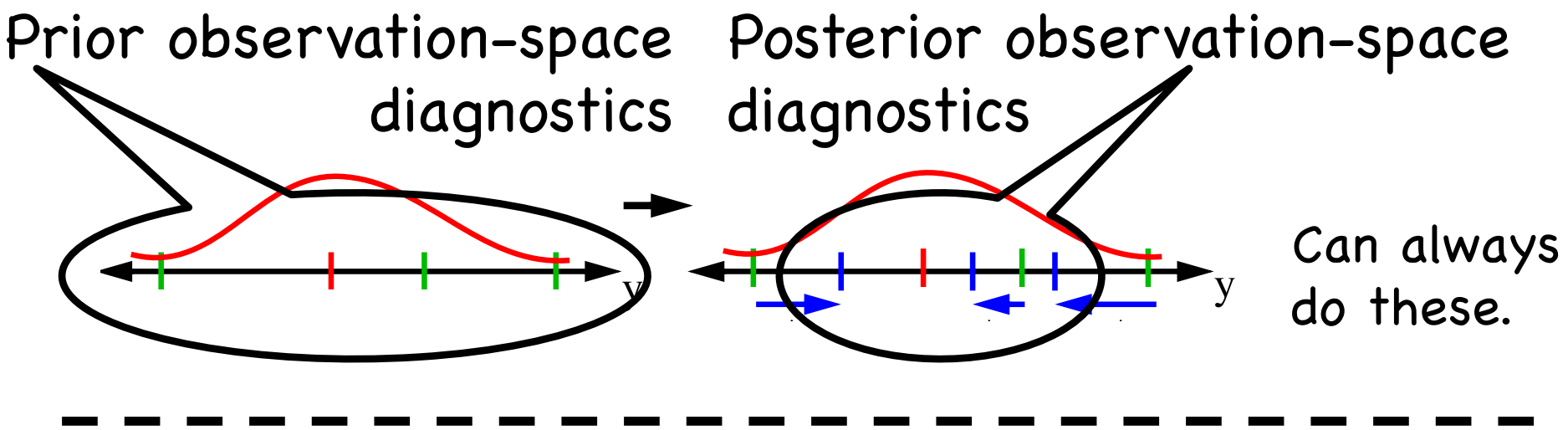
Some DART tools for STATE-space diagnostics:

routine	purpose
plot_bins.m	Plots the rank histograms (usually at a location)
plot_correl.m	plot the spatial correlation of the ensemble against a single location
plot_ens_time_series.m	Plots the evolution of the ensemble (1 location, 1 variable)
plot_total_error.m	Plots the RMSE of all variables, all locations* (if Truth is available)
plot_ens_err_spread.m	Plots the evolution of the ensemble error and spread



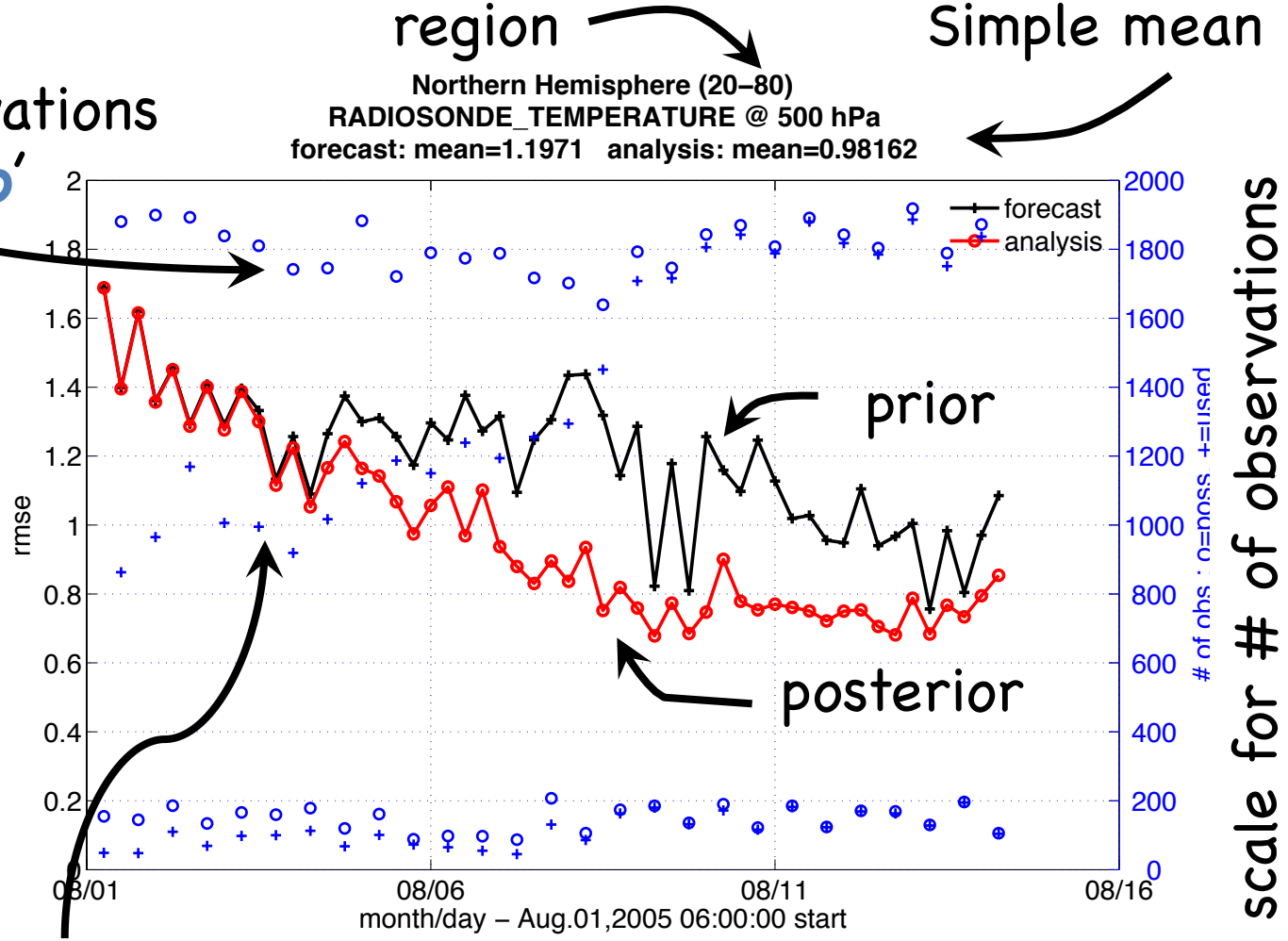
Some of these aggregate over a region,
some calculate metrics for each level,
some only make sense if you know the Truth …

Some DART tools for OBSERVATION-space diagnostics:



routine	purpose
obs_diag.f90	Calculates diagnostics from observations, writes out a netCDF file.
plot_evolution.m	plot the time-evolution of the diagnostic
plot_profile.m	plot a time-averaged vertical profile
plot_rank_histogram.m	plot rank histograms (can also be done with ncview)
obs_seq_to_netcdf.f90	creates a netCDF file of the observations
link_obs.m	Creates graphics to explore locations/values/QC etc.

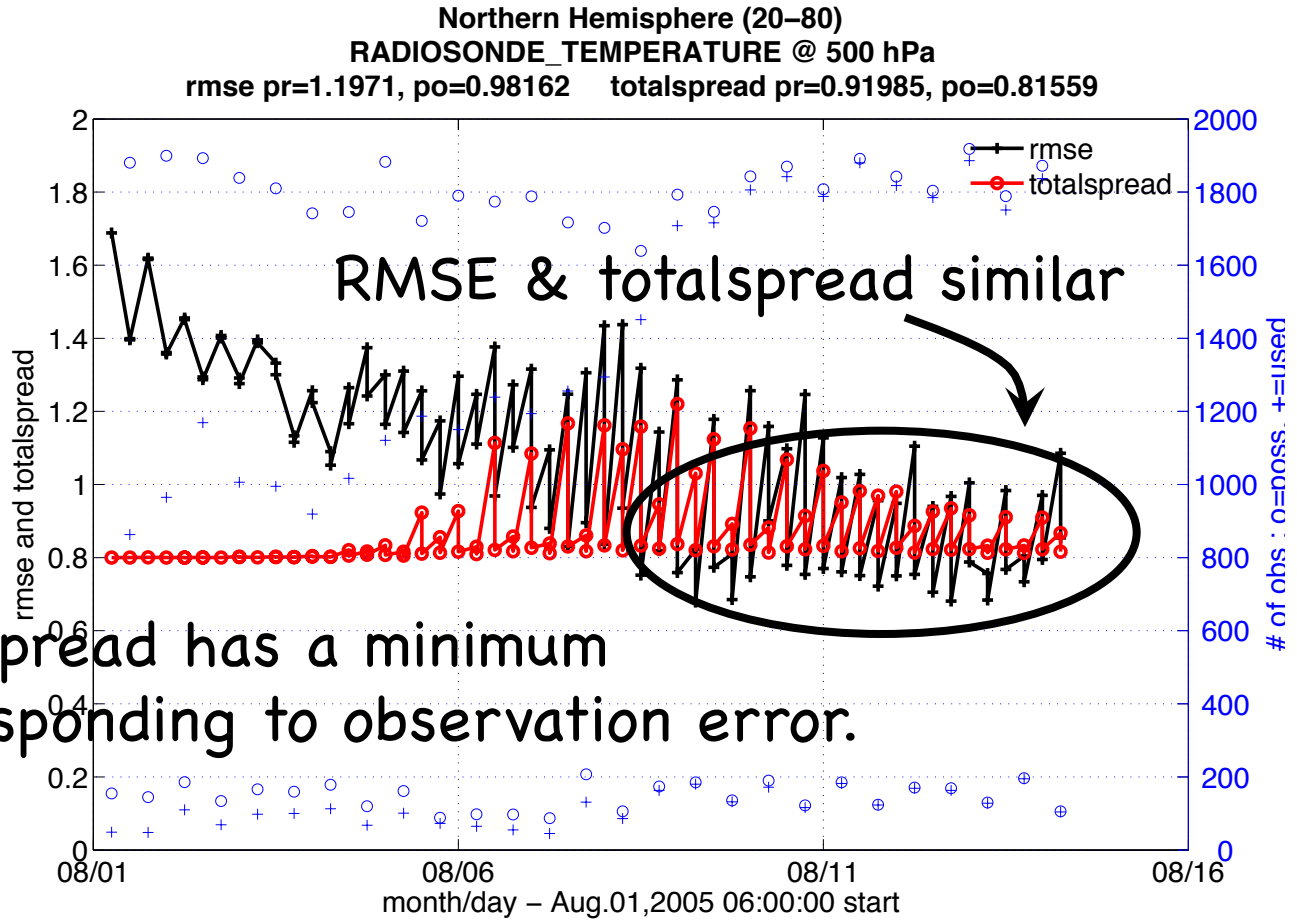
an example of plot_evolution



Can plot:
RMSE
Bias
Spread
Totalspread
Ens mean
QC value

This experiment started from nearly identical conditions and the model dynamics caused it to diverge over time.

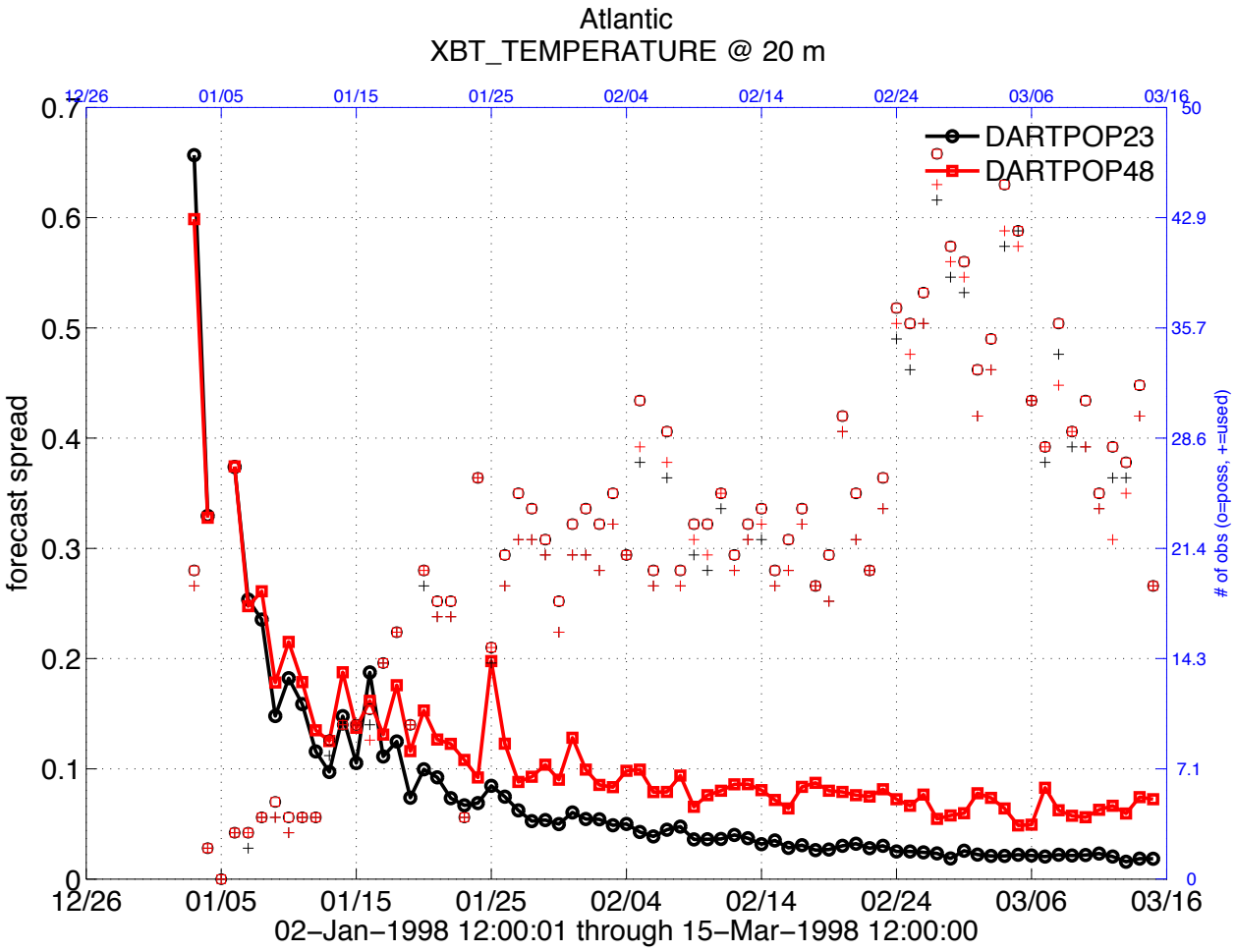
evolution of two quantities



Totalspread has a minimum corresponding to observation error.

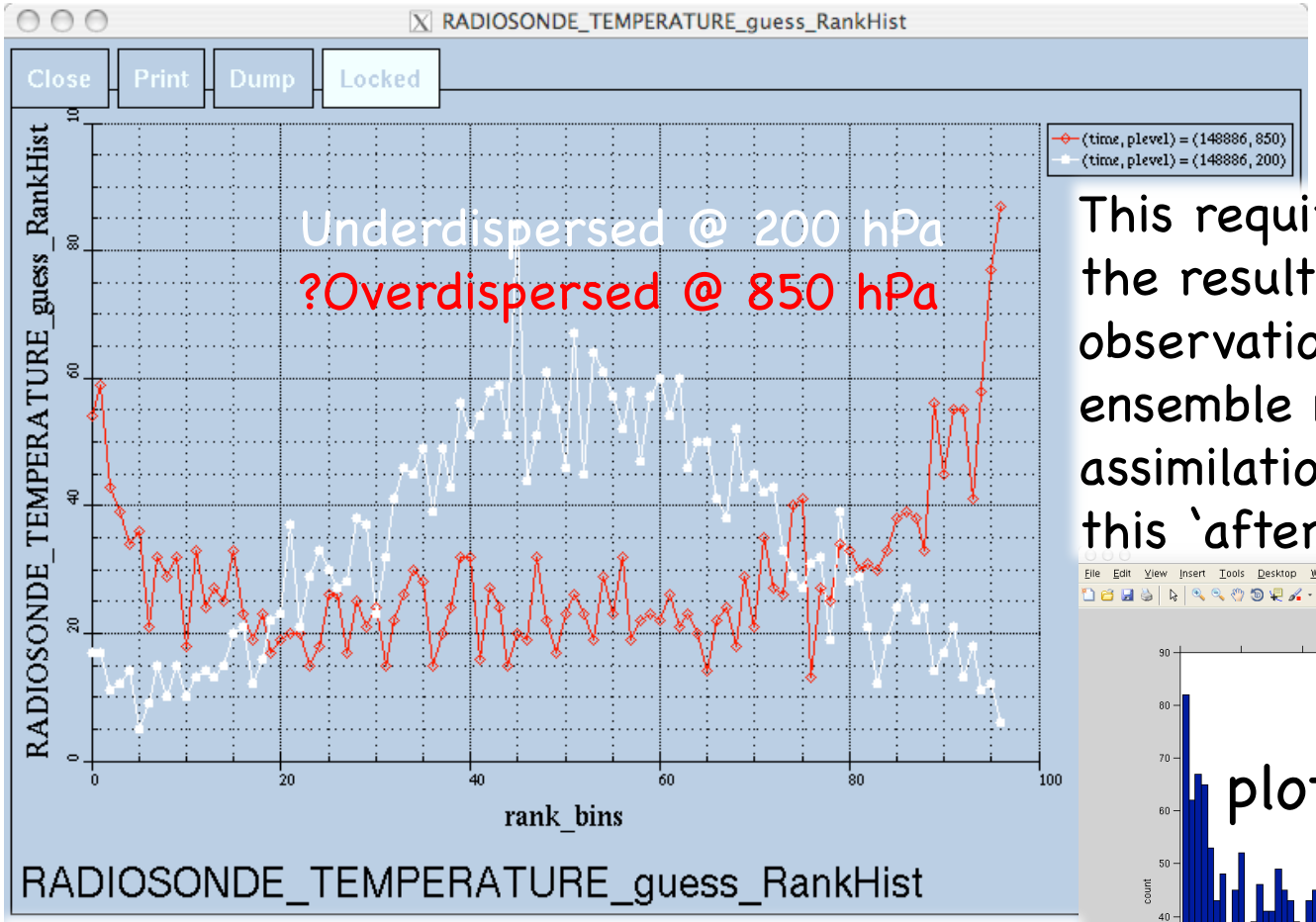
Prior and posterior plotted for each timestep … sawtooth.

compare multiple experiments

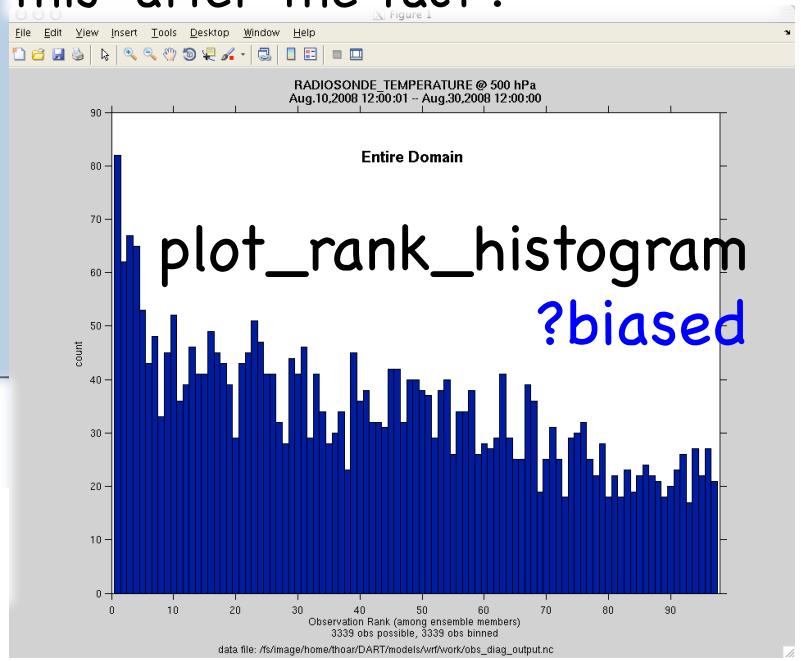


DART also has the ability to calculate the metrics for 'trusted' observations, i.e. even when the observation is rejected.

observation-space rank histograms

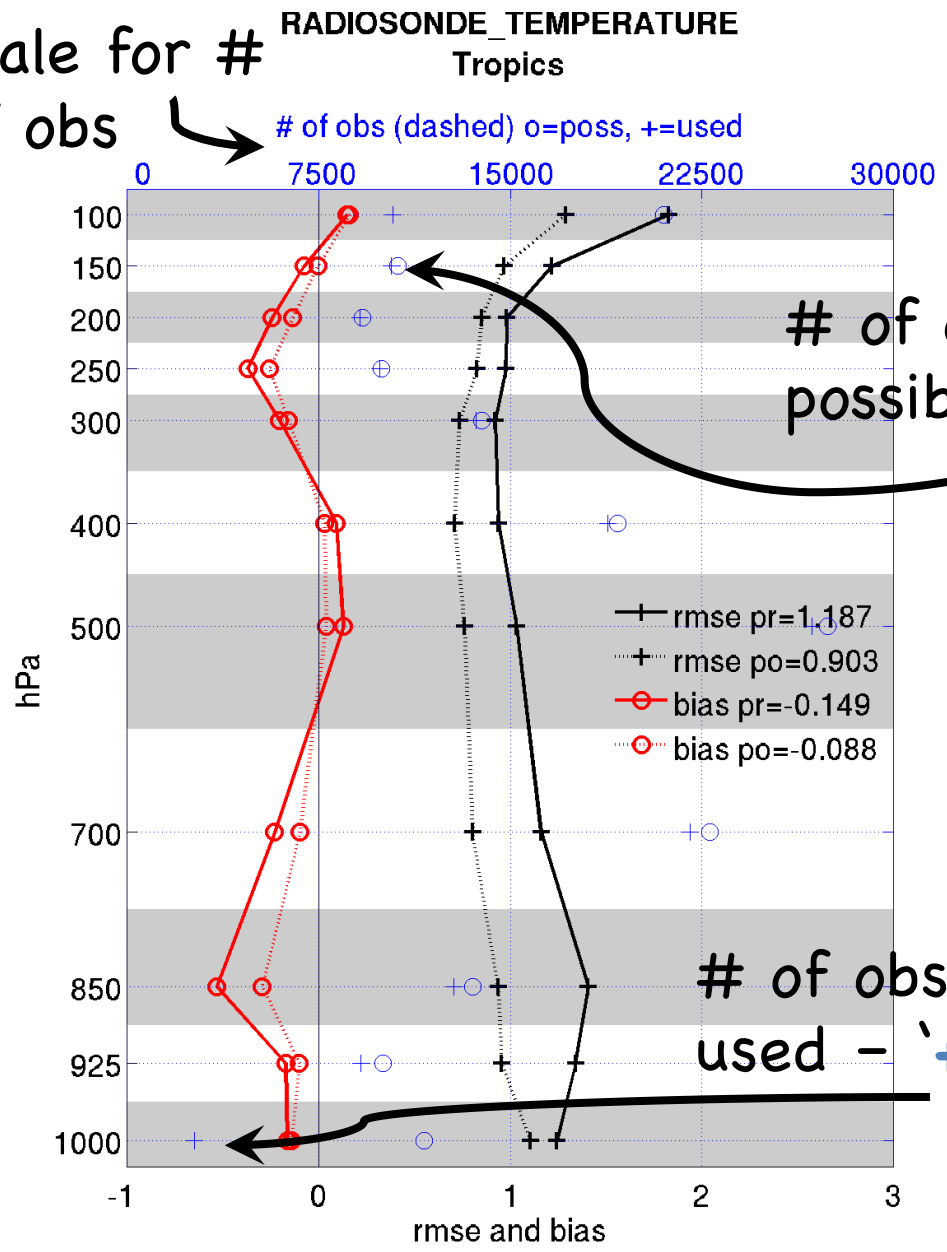


This requires the user to save the results of the forward observation operator for all ensemble members during the assimilation. Cannot calculate this 'after the fact'.



ncview can plot rank histograms right from obs_diag.f90 output. Can plot several levels on same axis.

an example of plot_vertical



scale for # of obs

Shaded areas correspond to vertical aggregation.

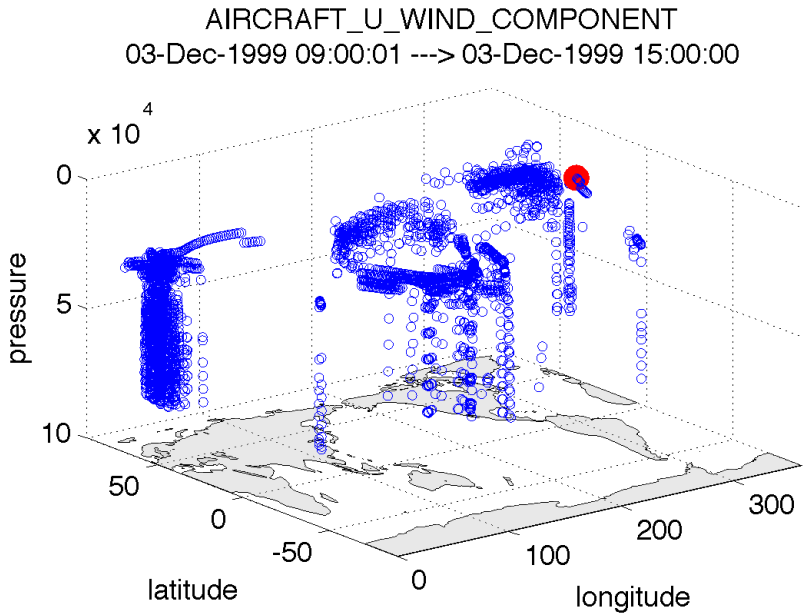
# of observations possible – 'o'

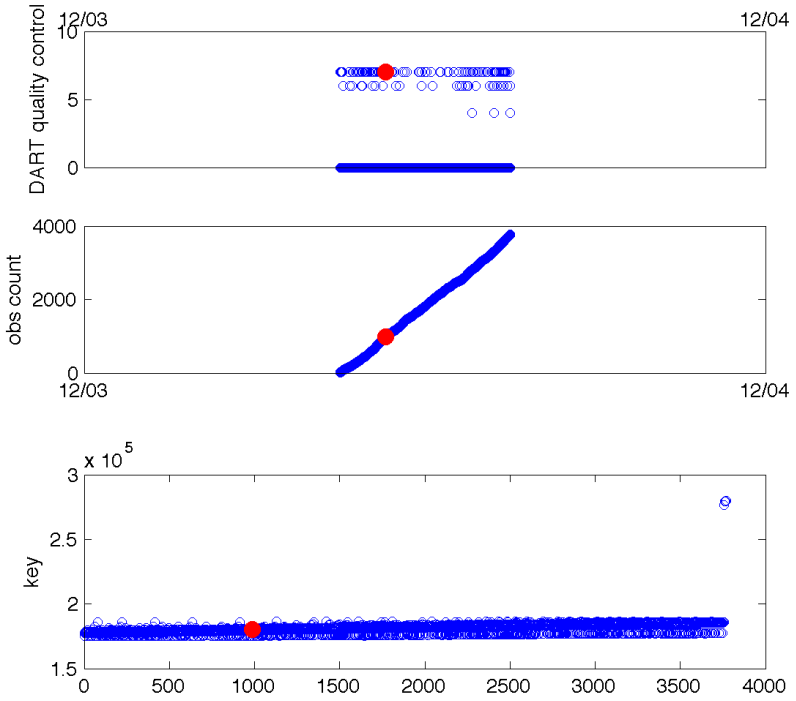
Important not to consider the 'spin up' period.

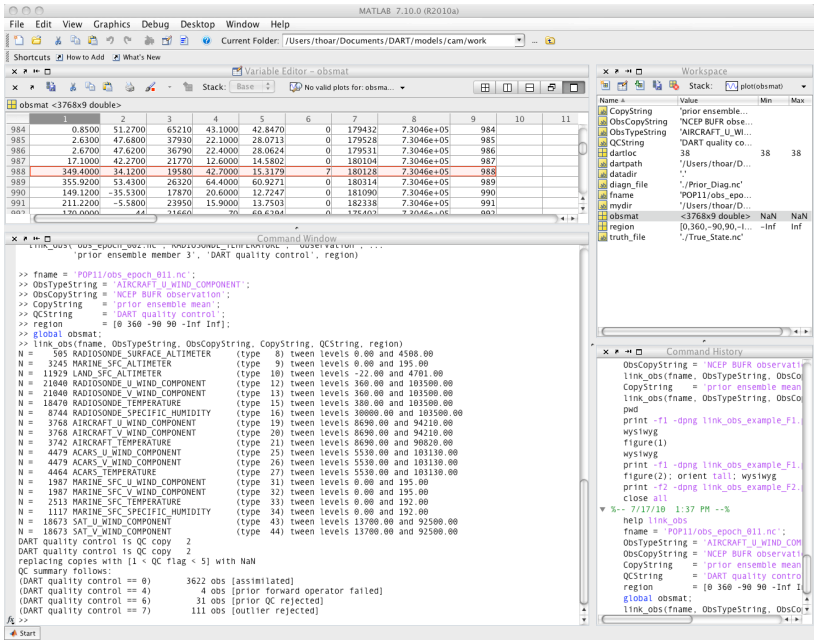
Priors are solid lines, posteriors are dashed.

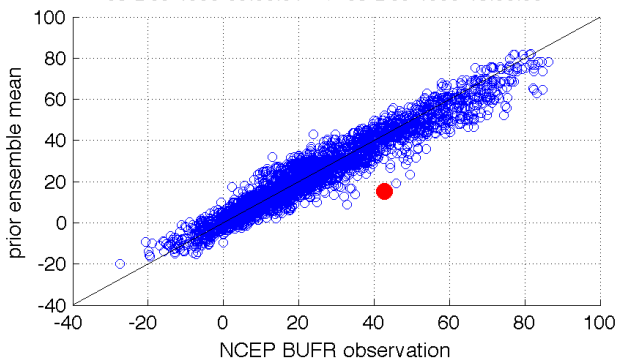
# of observations used – '+'

an example of link_obs

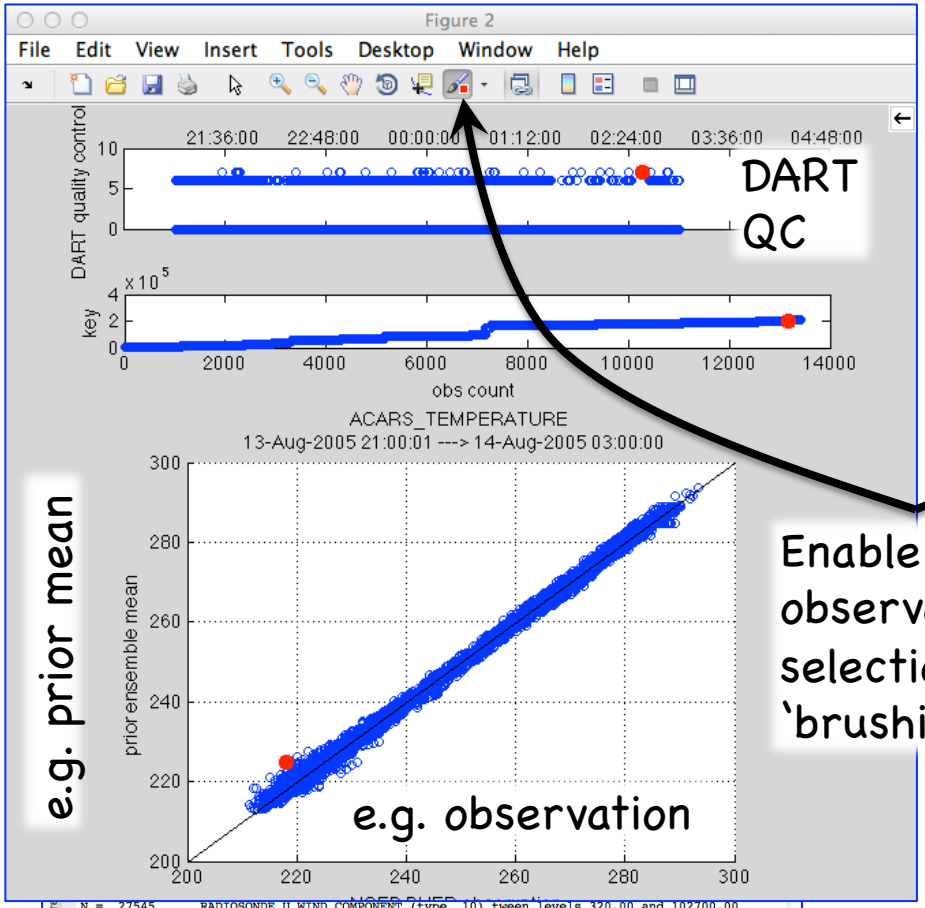


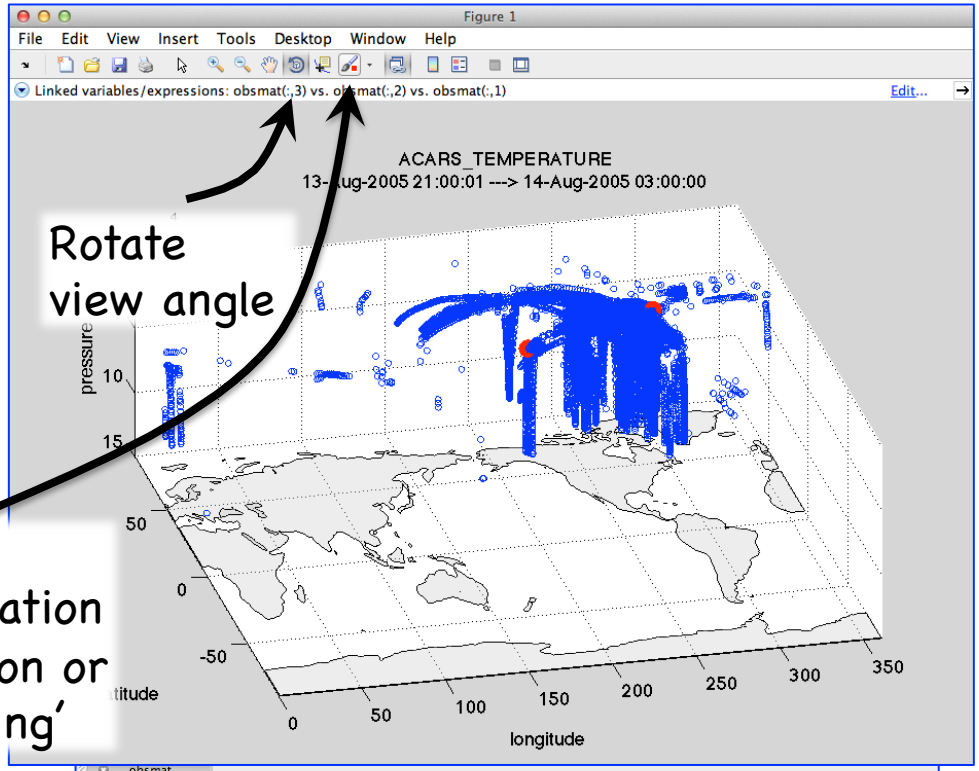






an example of link_obs






```
N =  27545      RADIOSONDE_U_WIND_COMPONENT (type  10) tween levels 320.00 and 102700.00
N =  27545      RADIOSONDE_V_WIND_COMPONENT (type  11) tween levels 320.00 and 102700.00
N =  22580         RADIOSONDE_TEMPERATURE (type  14) tween levels 360.00 and 102700.00
N =  11939  RADIOSONDE_SPECIFIC_HUMIDITY (type  15) tween levels 30000.00 and 102700.00
N =   6809       AIRCRAFT_U_WIND_COMPONENT (type  21) tween levels 17870.00 and 89150.00
N =   6809       AIRCRAFT_V_WIND_COMPONENT (type  22) tween levels 17870.00 and 89150.00
N =   4562          AIRCRAFT_TEMPERATURE (type  23) tween levels 17870.00 and 81200.00
N =  13465          ACARS_U_WIND_COMPONENT (type  25) tween levels 17870.00 and 101800.00
N =  13465          ACARS_V_WIND_COMPONENT (type  26) tween levels 17870.00 and 101800.00
N =  13409             ACARS_TEMPERATURE (type  27) tween levels 17870.00 and 101800.00
N =   2730    MARINE_SFC_U_WIND_COMPONENT (type  29) tween levels 0.00 and 195.00
N =   2730    MARINE_SFC_V_WIND_COMPONENT (type  30) tween levels 0.00 and 195.00
N =   2972       MARINE_SFC_TEMPERATURE (type  31) tween levels 0.00 and 195.00
N =   1315  MARINE_SFC_SPECIFIC_HUMIDITY (type  32) tween levels 0.00 and 192.00
N =  17813           SAT_U_WIND_COMPONENT (type  39) tween levels 13340.00 and 99280.00
N =  17813           SAT_V_WIND_COMPONENT (type  40) tween levels 13340.00 and 99280.00
DART quality control is QC copy   2
DART quality control is QC copy   2
replacing copies with [1 < QC flag < 5] with NaN
QC summary follows:
(DART quality control == 0)       10747 obs [assimilated]
(DART quality control == 6)        2623 obs [prior QC rejected]
(DART quality control == 7)          39 obs [outlier rejected]
>>
```

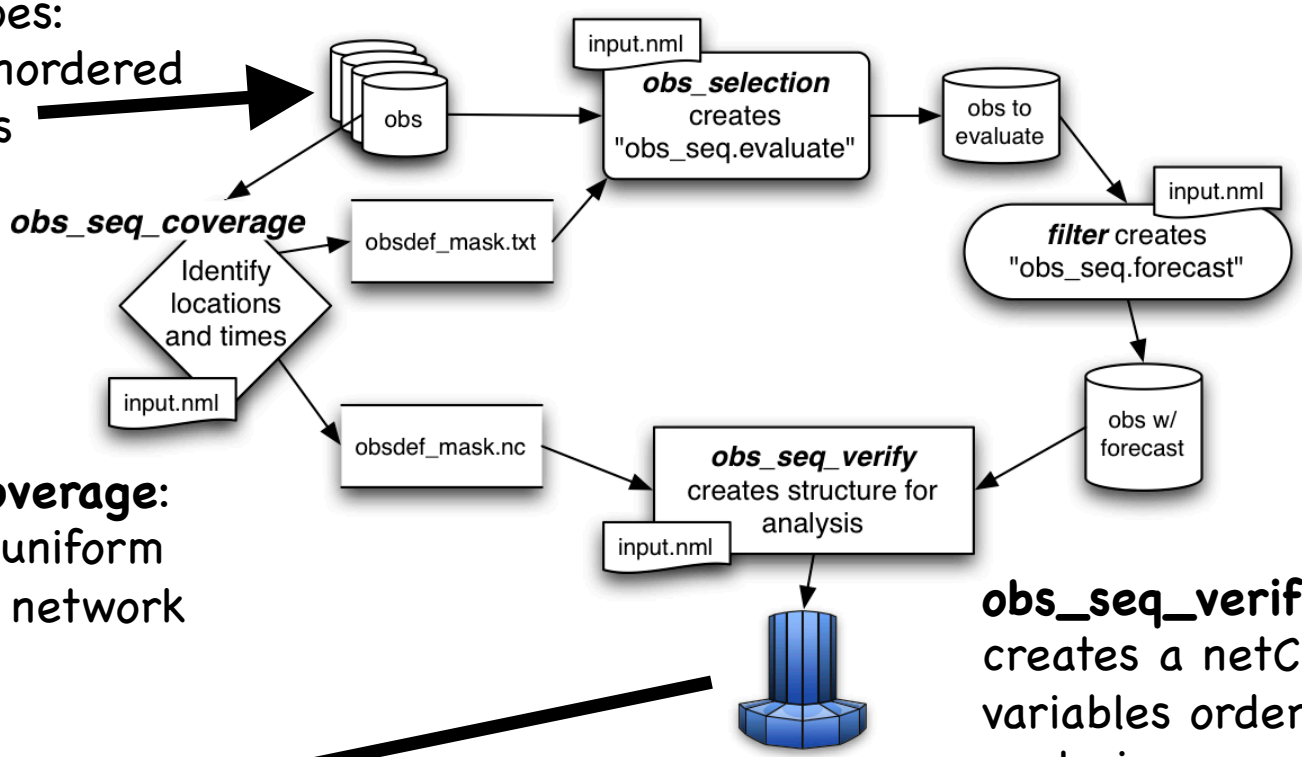
obsmat

obsmat <13409x9 double>

	1	2	3	4	5	6	7	8	9	1
13174	320	54	25000	222.1600	223.7053	0	204376	7.3254e+05	13174	
13175	211.6700	61.2700	42850	253.8600	252.6254	0	204379	7.3254e+05	13175	
13176	234.5200	54.0600	22730	219.1600	218.6082	0	204382	7.3254e+05	13176	
13177	205.9500	22.5300	22730	227.3600	225.8189	0	204385	7.3254e+05	13177	
13178	186.7100	57.4400	26190	224.1600	224.9040	0	204388	7.3254e+05	13178	
13179	211.8000	61.2700	41380	251.9600	250.5536	0	204452	7.3254e+05	13179	
13180	259.9700	25.1700	53590	271.3600	271.4563	0	204455	7.3254e+05	13180	
13181	298.3900	59.7300	23840	218.1600	224.6808	7	204458	7.3254e+05	13181	
13182	231.8100	58.4200	20650	214.8600	214.1854	0	204461	7.3254e+05	13182	
13183	210.5900	61.0300	73270	282.2600	281.8255	0	204464	7.3254e+05	13183	
13184	234.0900	36.7500	24990	224.8600	225.0371	0	204467	7.3254e+05	13184	
13185	176.6200	54.6500	28740	235.0600	234.3654	0	204470	7.3254e+05	13185	
13186	198.1800	19.6200	30090	242.3600	241.1061	0	204473	7.3254e+05	13186	
13187	238.8800	59.2600	22730	217.6600	216.6183	0	204476	7.3254e+05	13187	
13188	320.0800	53.0100	23840	222.1600	221.6454	0	204542	7.3254e+05	13188	
13189	278.1600	13.9100	21660	224.1600	223.2794	0	204545	7.3254e+05	13189	
13190	259.9600	25.2600	56760	273.8600	273.8386	0	204548	7.3254e+05	13190	
13191	101.1900	68.2400	24070	231.1600	231.3765	0	204551	7.3254e+05	13191	

Observation processing for forecast metrics:

Anything goes:
unsorted, unordered
observations



obs_seq_coverage:
determines uniform
observation network

obs_seq_verify:
creates a netCDF file with
variables ordered to make
analysis easy

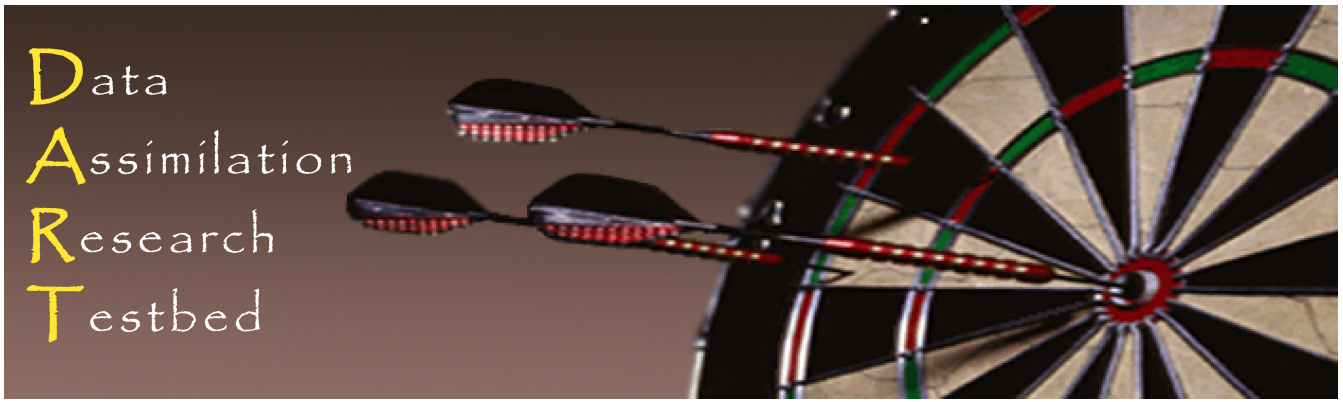
VariableX:

```
(analysisT, station, level, copy, ensemble, forecast_lead)
     |         |       |      |       |          |
     |         |       |      |       |          +- forecast length : 0,3,6,9,12,...
     |         |       |      |       +------------ ensemble member index
     |         |       |      +-------------------- obs value, prior, obs_err
     |         |       +--------------------------- vertical level index
     |         +----------------------------------- (horizontal) station index
     +--------------------------------------------- analysis time/date
```

For more information:

CAM GITM ROMS wrfHydro WACCM WRF

CLM AM2



POP BGRID

COAMPS NOAH

www.image.ucar.edu/DAReS/DART

dart@ucar.edu

MPAS_ATM MITgcm_ocean COAMPS_nest TIEGCM SQG NAAPS MPAS_OCN PBL_1d PE2LYR NCOMMAS CABLE