

Comparing CAM-FV and CAM-SE in a 'perfect model' data assimilation environment

Kevin Raeder, Jeff Anderson,
Tim Hoar, Nancy Collins

NCAR/CISL/IMAGE/DAReS

- The Data Assimilation Research Testbed has assimilated many kinds of observations into many versions of CAM (eulerian and finite volume 'FV')
- Results are comparable to operational forecasting centers.
- Extensions into CESM:
 - + POP and CLM
 - + Assimilation into a coupled model: CAM+POP+CLM
 - + WACCM (Pedatella & Liu in WAWG)
 - + CAM-Chem (Barré in CCWG)
- Now extended to spectral element CAM ('SE')

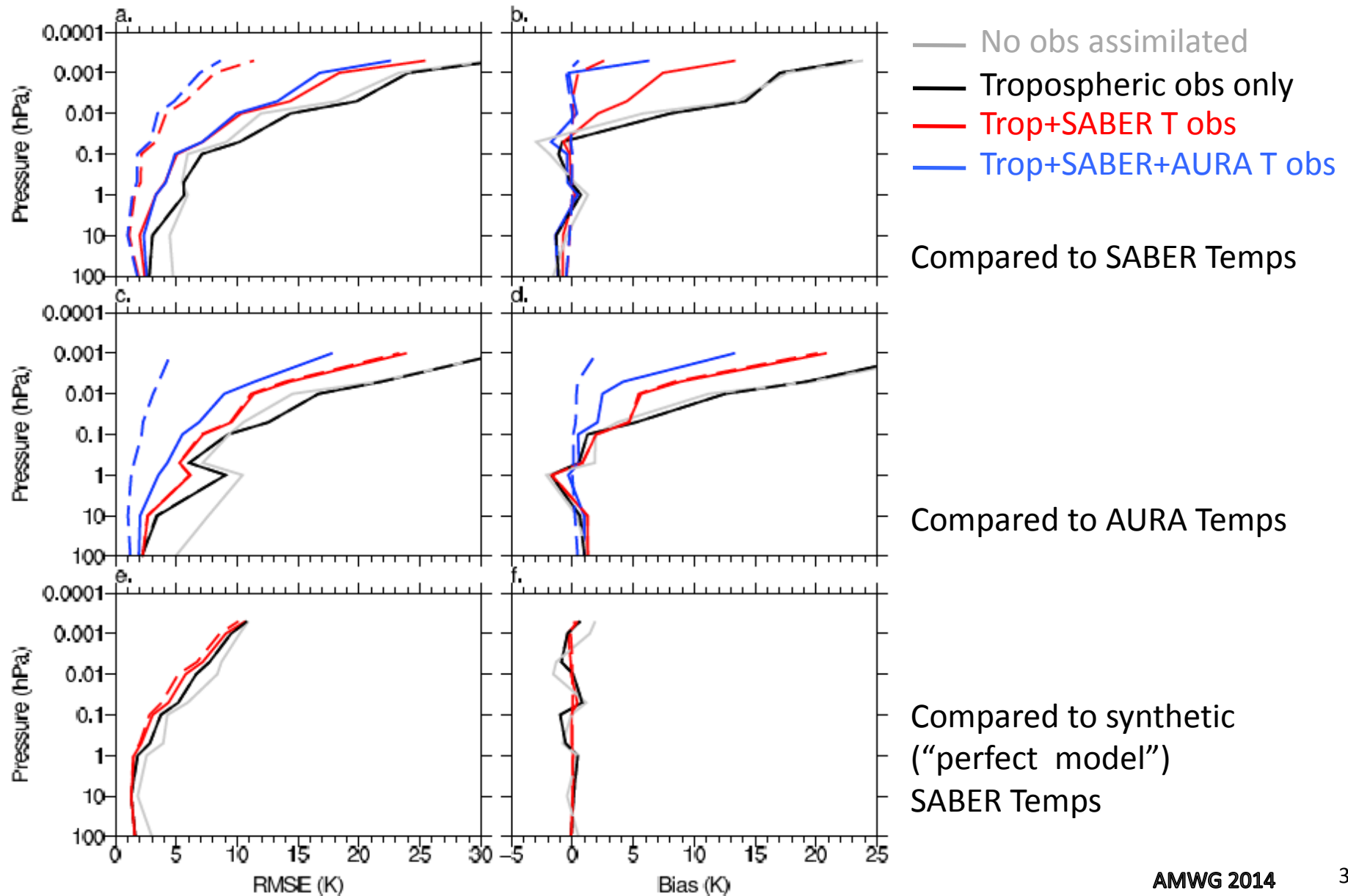
1-slide tangent: WACCM+DART results

Then compare CAM-SE against CAM-FV in a 'perfect model' context; the first validation step of a data assimilation implementation.

WACCM(FV)+DART: SABER and AURA



Pedatella, et al. "Ensemble Data Assimilation in the Whole Atmosphere Community Climate Model" JGR, in press



Perfect Model

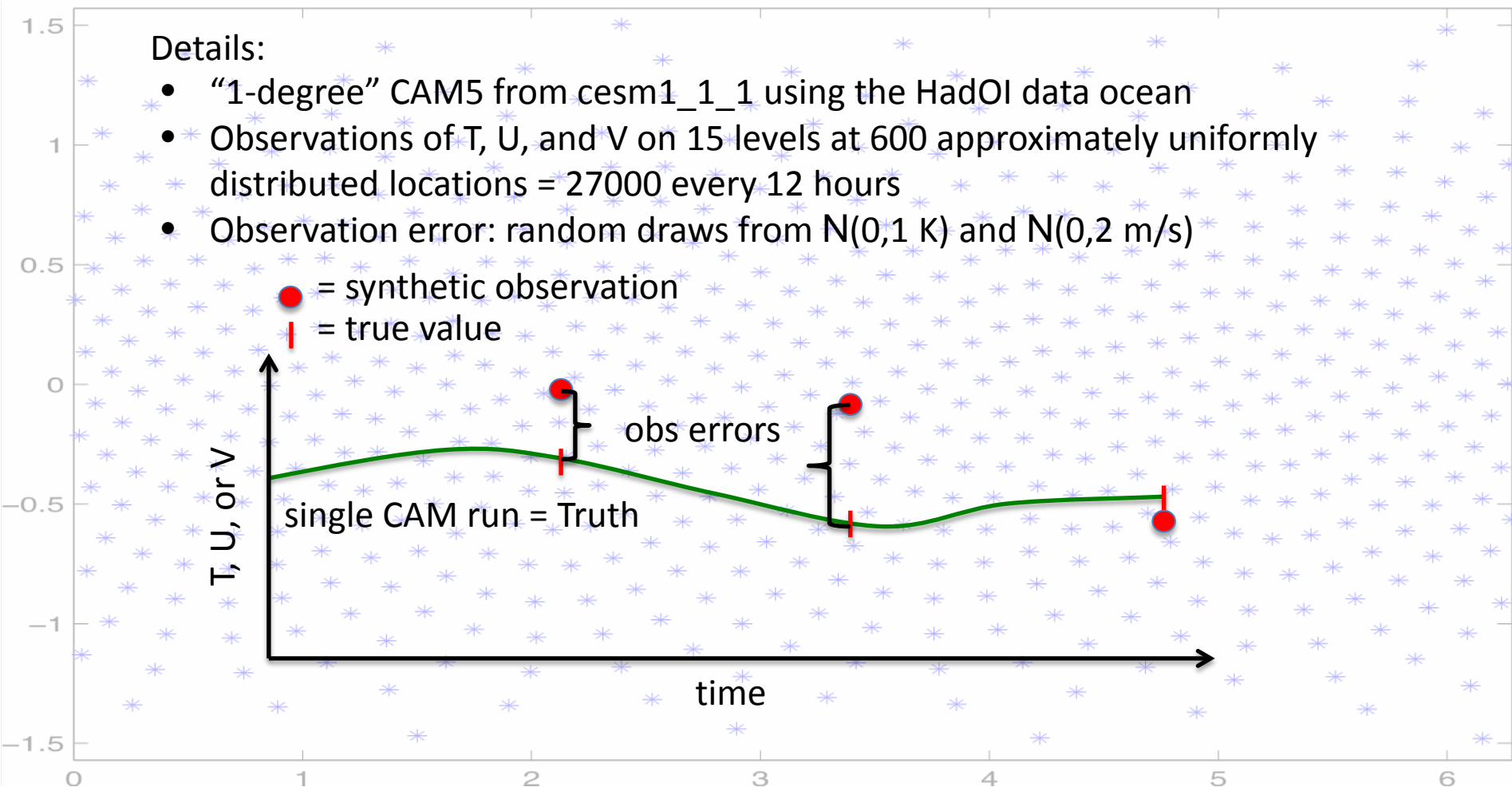
Free run of the model is used as the Truth.

Observations of the evolving model state are taken periodically.

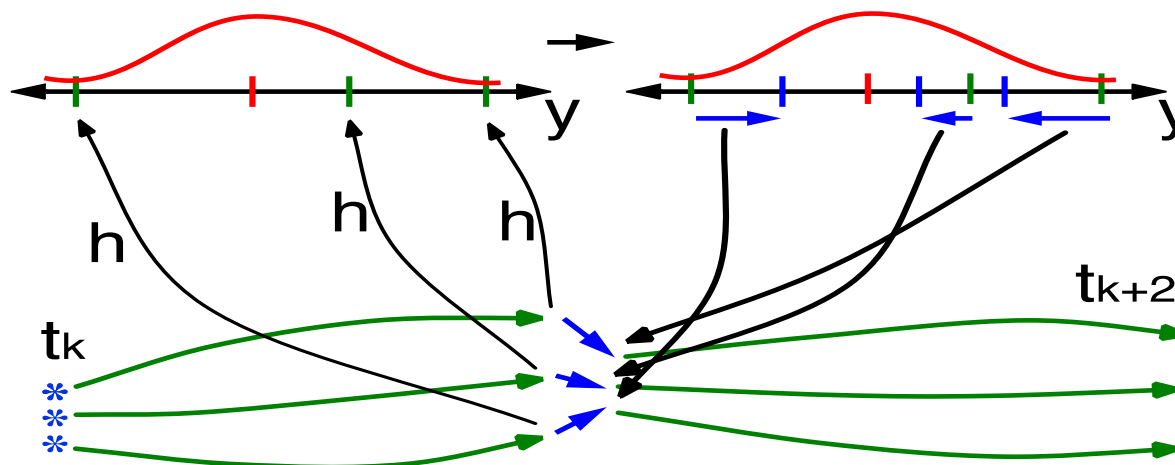
Observational error is added to the observed values to make them realistic.

Details:

- “1-degree” CAM5 from cesm1_1_1 using the HadOI data ocean
- Observations of T, U, and V on 15 levels at 600 approximately uniformly distributed locations = 27000 every 12 hours
- Observation error: random draws from $N(0,1 \text{ K})$ and $N(0,2 \text{ m/s})$



Ensemble Kalman filter algorithm requires running an **ensemble** of equally likely **forecasts**, then using statistics of the ensemble and **observations** to allow the observations to guide the ensemble to a better description of the atmosphere.



Ensemble spread = uncertainty of the model state.

It grows during forecast, shrinks as information is added during assimilation.

Neglected errors -> ensemble spread is too small (over confident).

This can be fixed with “adaptive inflation”, which increases ensemble spread but retains the mean.

Observations more than $3 \times \text{spread}$ away from ensemble mean are ignored as being erroneous.

“Identical Twin” Experiments



Will the ensemble track the observations (Truth + obs error) if the Truth model = the DA model = CAM-SE?

Initial ensemble has a tiny spread ($O(\text{round-off})$ in T) centered around the True State.

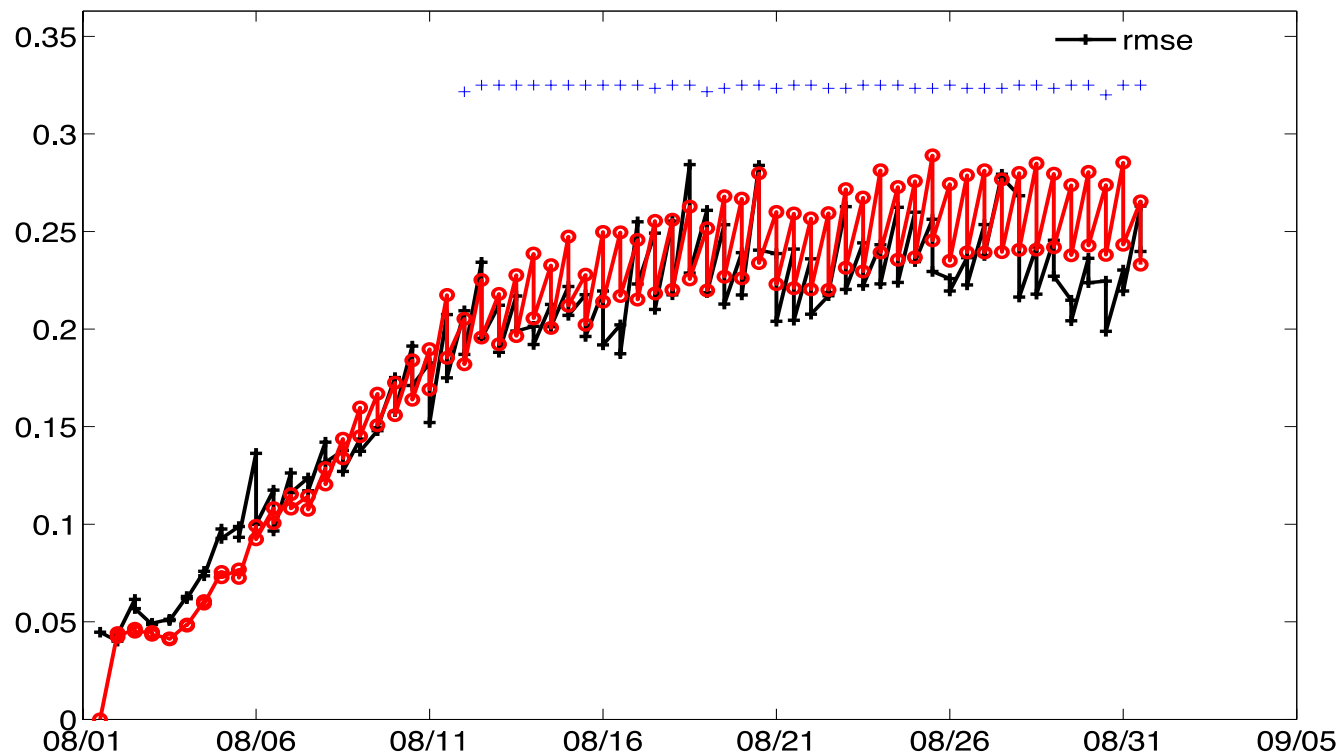
No inflation (shouldn't be needed).

“Identical Twin” Experiments



Will the ensemble track the observations (Truth + obs error) if the Truth model = the DA model = CAM-SE?

Initial ensemble has a tiny spread ($O(\text{round-off})$ in T)
centered around the True State.
No inflation.



of obs : 0=pos, +=used

**Southern Hemisphere
RADIOSONDE TEMPERATURE @ 500 hPa**



“Fraternal Twin” Experiments



Will the ensemble of CAM-**SE** track the observations of CAM-**FV**?
And vice versa.

A harder test: still start with tiny spread, but **wrong model state**.

Southern Hemisphere
RADIOSONDE_TEMPERATURE @ 500 hPa

—○— FV flw in
—■— SE flw in

forecast rmse

01-Aug-2005 06:00:01 through 01-Sep-2005 18:00:00

data file: /global/work/forecast/rpt/V-SE_perfect/V-RUU flwml_obs_dmg.nc
data file: /global/work/forecast/rpt/V-SE_perfect/V-RUU flwml_obs_dmg.nc

“Fraternal Twin” Experiments



Will the ensemble of CAM-SE track the observations of CAM-FV?
And vice versa.

No, because the initial ensemble spread is so small that the assimilation ignores the large number of ‘outlier’ observations taken from the other model.

“Fraternal Twin” Experiments



Will the ensemble of CAM-SE track the observations of CAM-FV?
And vice versa.

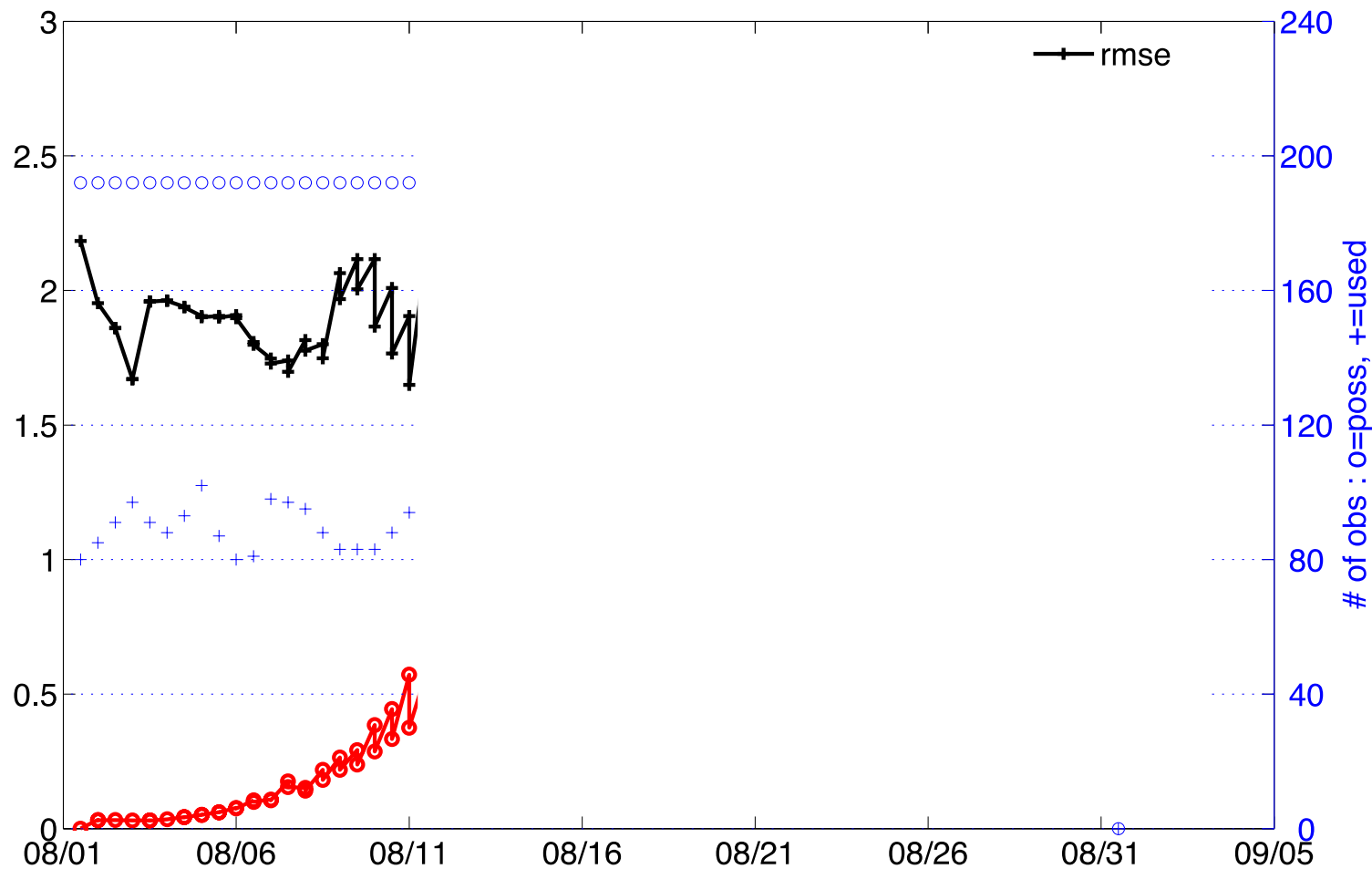
No, because the initial ensemble spread is so small that the assimilation ignores the large number of ‘outlier’ observations taken from the other model.

What if we give the assimilation a fighting chance by turning on the adaptive inflation? That may allow the spread to increase, which will allow more observations to be assimilated.

Fraternal Twins with Inflation



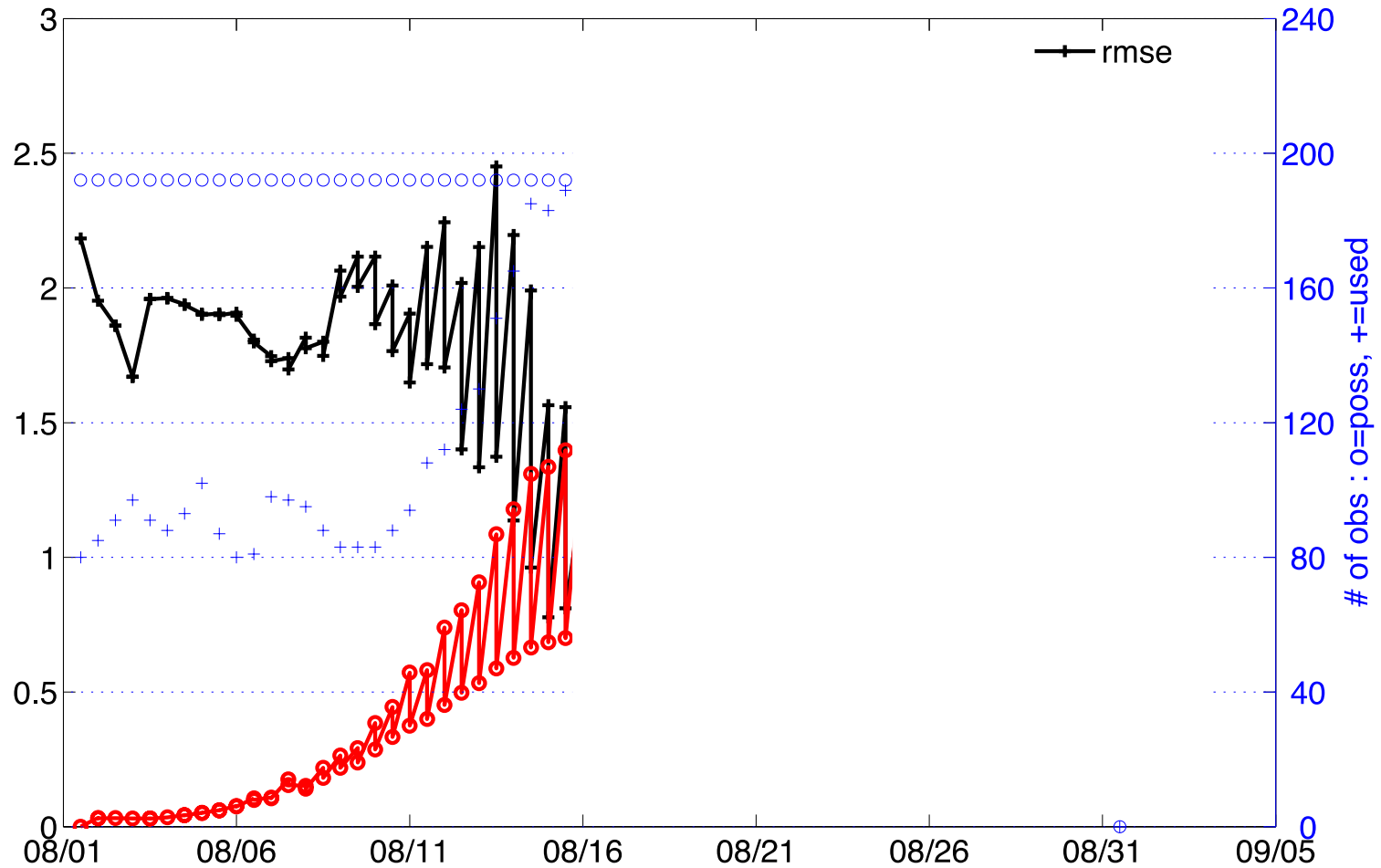
Obs from CAM-FV. Assimilating model is CAM-SE.



Fraternal Twins with Inflation



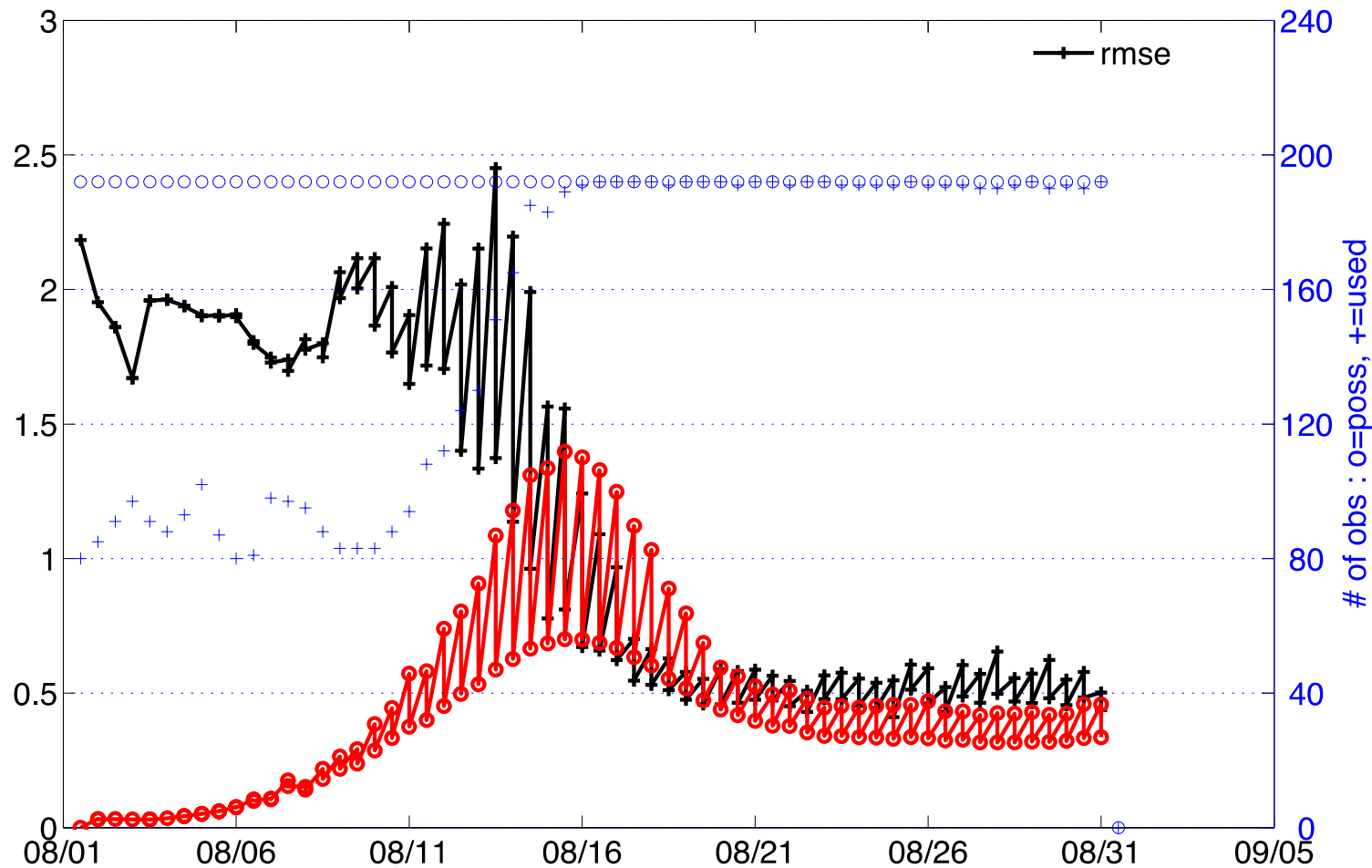
Obs from CAM-FV. Assimilating model is CAM-SE.



Fraternal Twins with Inflation

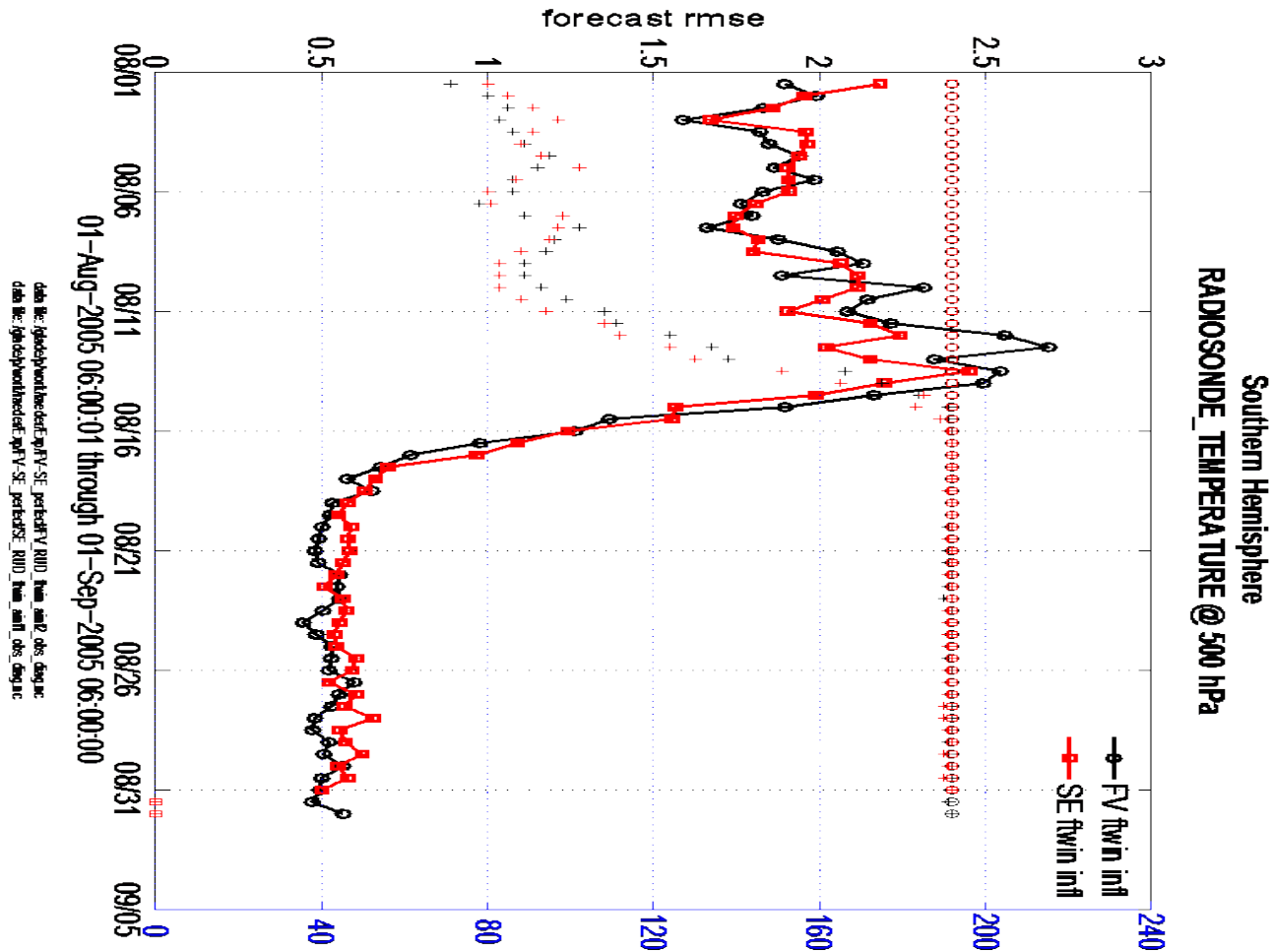


Obs from CAM-FV. Assimilating model is CAM-SE.

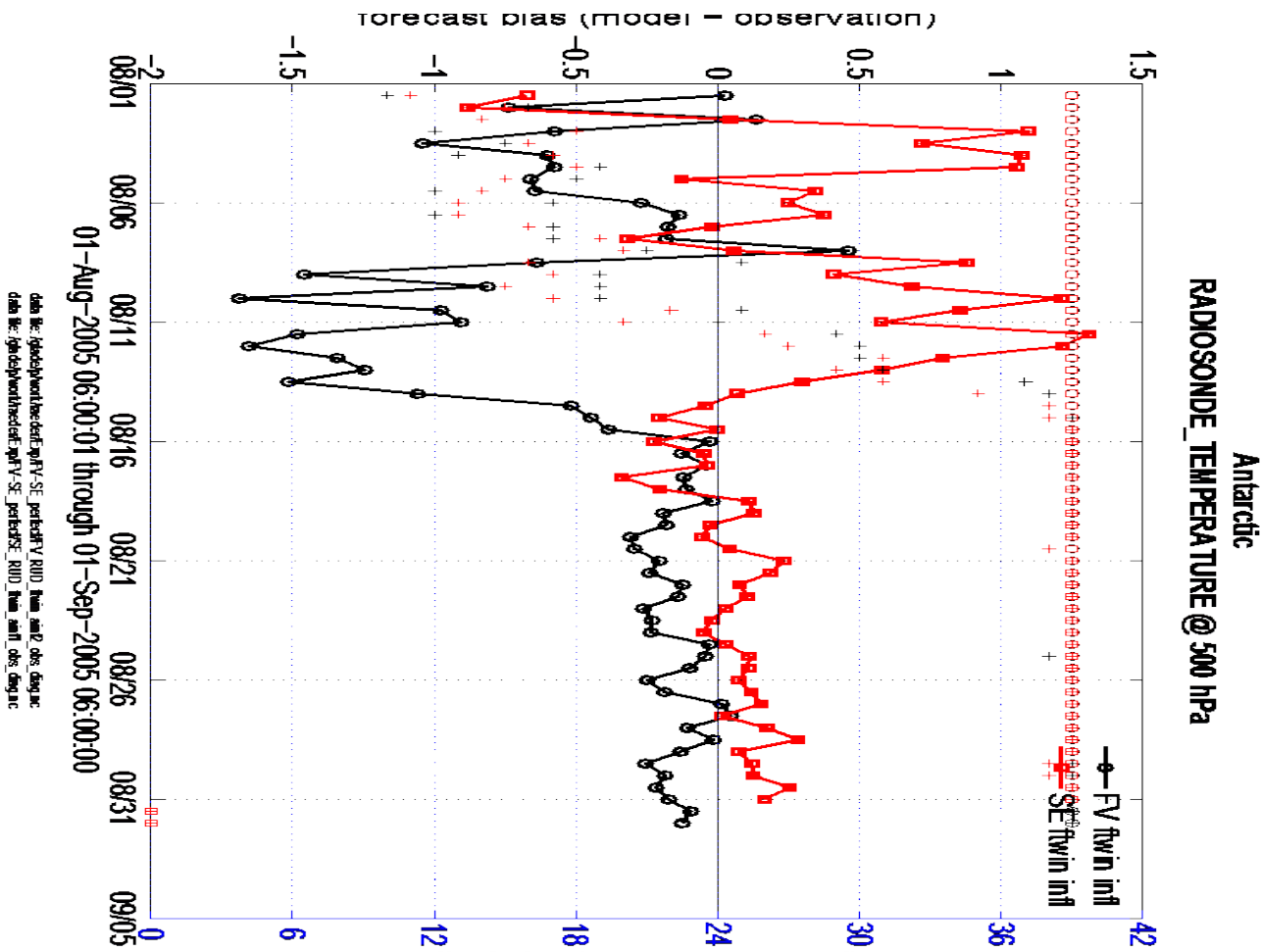


“SE ftwin infl” = Obs from CAM-FV. Assimilating model is CAM-SE.

“FV ftwin infl” = Obs from CAM-SE. Assimilating model is CAM-FV.



“FV ftwin infl” = Obs from CAM-SE. Assimilating model is CAM-FV.



State Space Confirmation

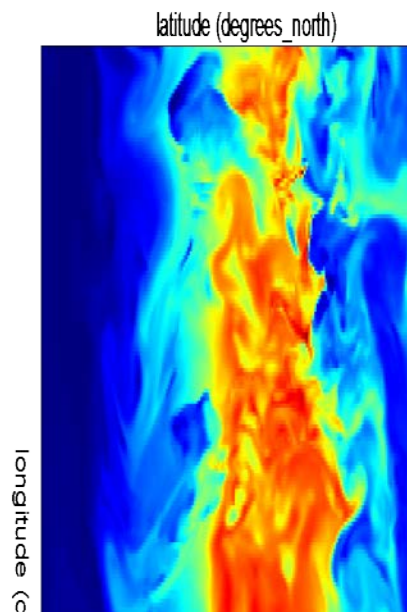


Compare 2 assimilations which use the **same** set of observations:

- 1) Identical twin using CAM-FV for obs and assimilation
- 2) Fraternal twin using CAM-FV for obs, but CAM-SE for assimilation

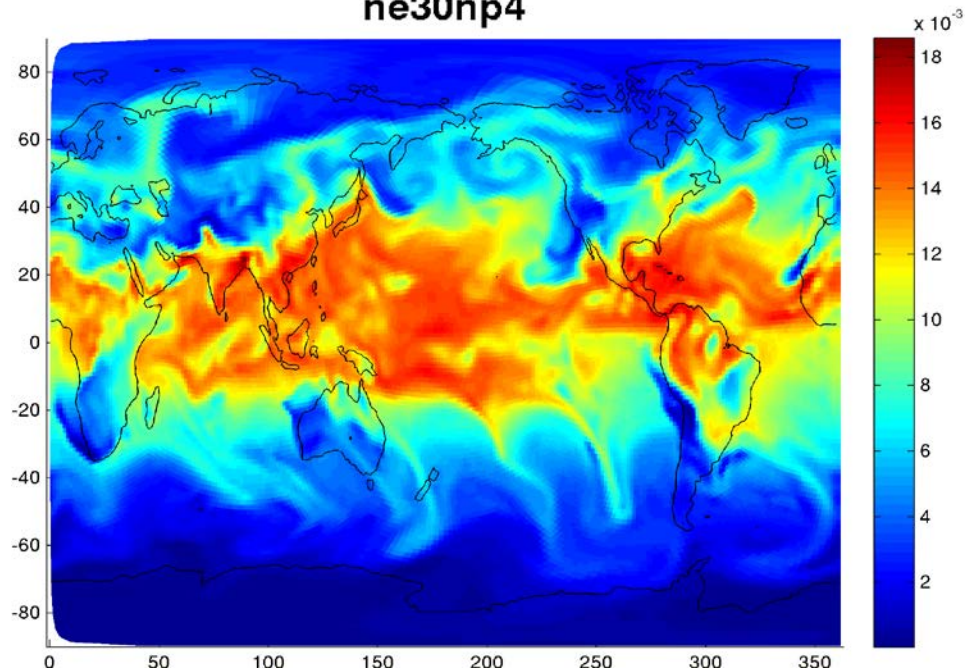
1)

Ensemble mean



2)

Ensemble mean
Q level 25
ne30np4



Q is not an observed variable, but has been brought to the observed model state.

Conclusions



- Data Assimilation with DART and CAM-SE passes the 'perfect model' tests.
- The adaptive inflation algorithm effectively permits even ensembles with small spread and wrong mean to be shifted to the correct model state.
- Assimilations with CAM-SE appear comparable to those with CAM-FV, although much more analysis could be done. Collaboration?
- This tool can identify some biases in a new model by comparison with known biases in an old model, without a long run of the new model.
- A new model can be searched for problem areas by direct comparison to observations.
- We're ready for a real observation test of CAM-SE+DART, including observations that are not model variables.

For more information:

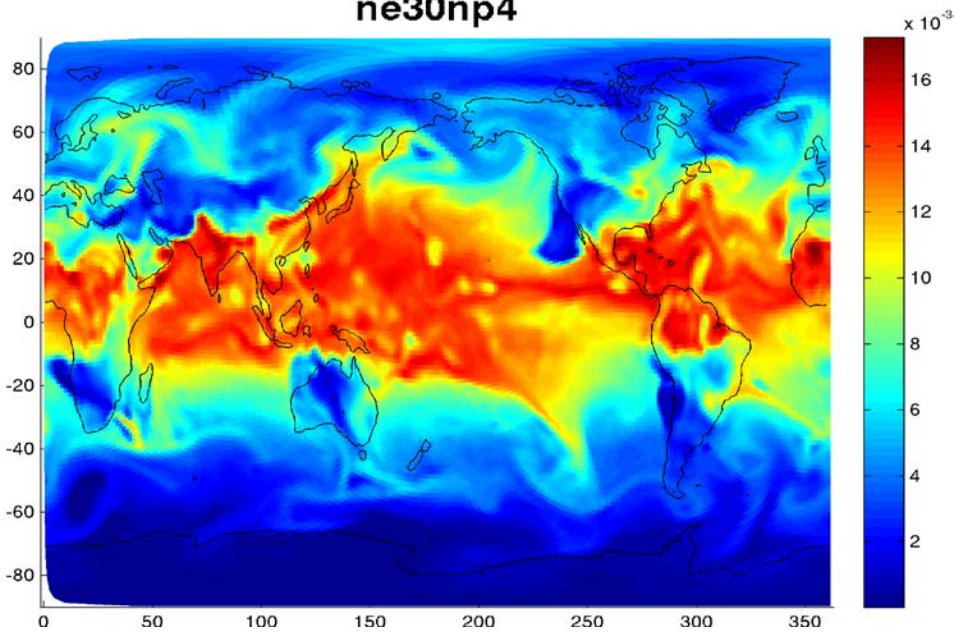
<http://www.image.ucar.edu/DARes/DART/>

(no facebook, no twitter)



5 nodes/instance (ensemble member) (usually 3, but sporadic memory problems)
~2200 core hours/ 12 hour advance and assimilate 27000 obs.
Wall clock; < 30 min/(forecast+assim cycle)
+ waiting between jobs
~200 Gb/ensemble restart set. Save infrequently
Up to 10 Gb/assimilation of DART output. So up to 0.5 Tb/month, user selectable.

**Q level 25
ne30np4**



**Q level 25
ne30np4**

