Predictability in a simulated global prediction system that assimilates only surface pressure observations

Jeffrey Anderson
NOAA/GFDL & NCAR/DAI

Special thanks to:
Bruce Wyman
Tim Hoar
Shaoqing Zhang

Data Assimilation Research Testbed (DART) update

Ensemble filter assimilation

Model description

Many assimilation system simulation experiments

Conclusions

The Data Assimilation Research Testbed (DART)

What is DART?

1. Allows combinations of assimilation algorithms, models, and observation sets
2. Diagnostic tools
3. Supports Data Assimilation R&D for NOAA/NCAR and external partners
   NOT for operational use or support

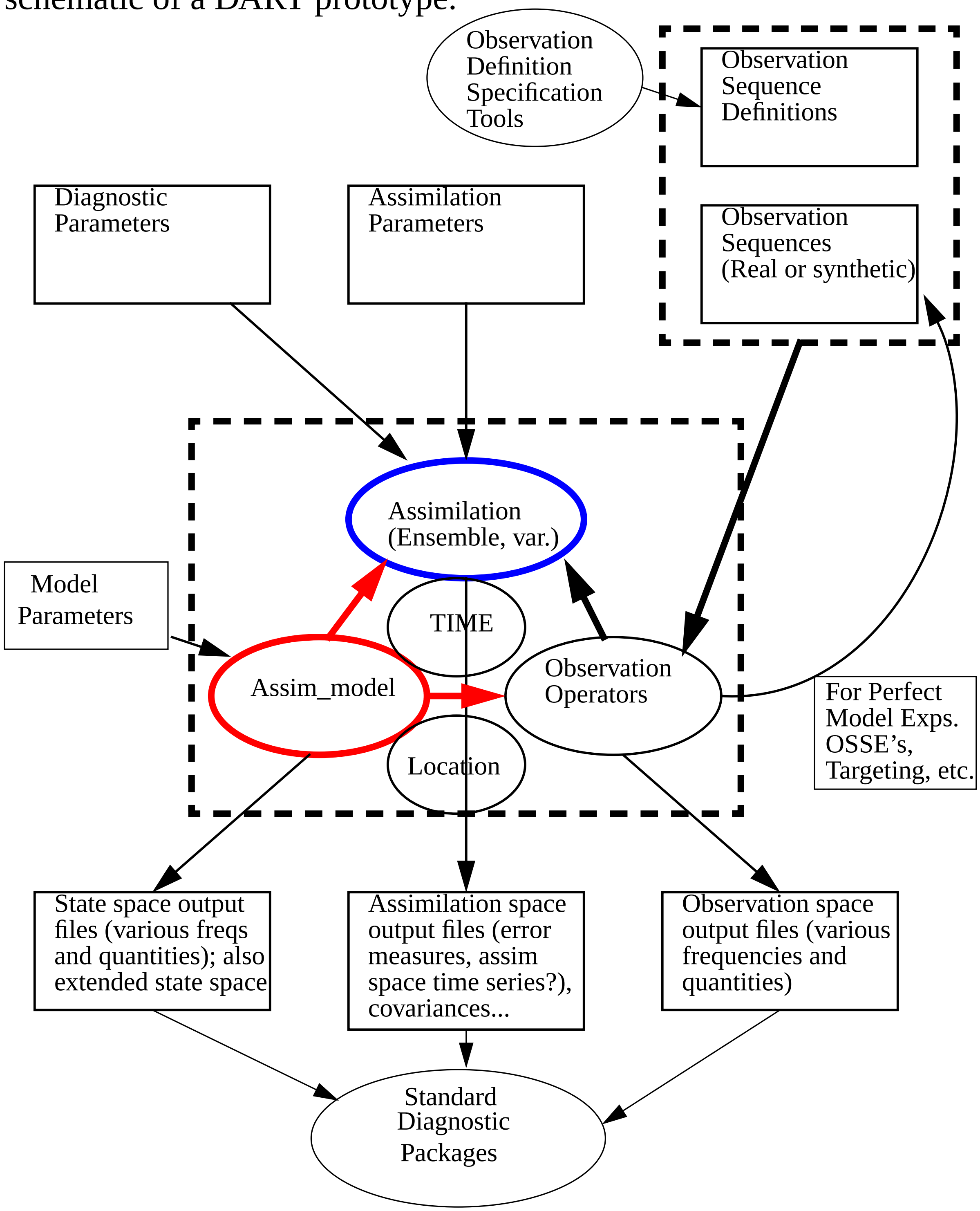
Status of DART

1. Basic framework implemented
2. Currently using GFDL FMS infrastructure
3. Switch to ESMF infrastructure when available
4. Primarily implementing ensemble (Kalman) filters
5. Variational for low-order models only
6. Plans MAY include a variational (4D-Var) capability

DART compliant models

1. GFDL FMS B-grid GCM incorporated and in use
2. Many low-order models available
3. WRF model in process of being incorporated
4. NCEP MRF being tested quasi-operationally in partial implementation
5. GFDL MOM ocean model partially incorporated in earlier version
6. Initial work on incorporating CAM

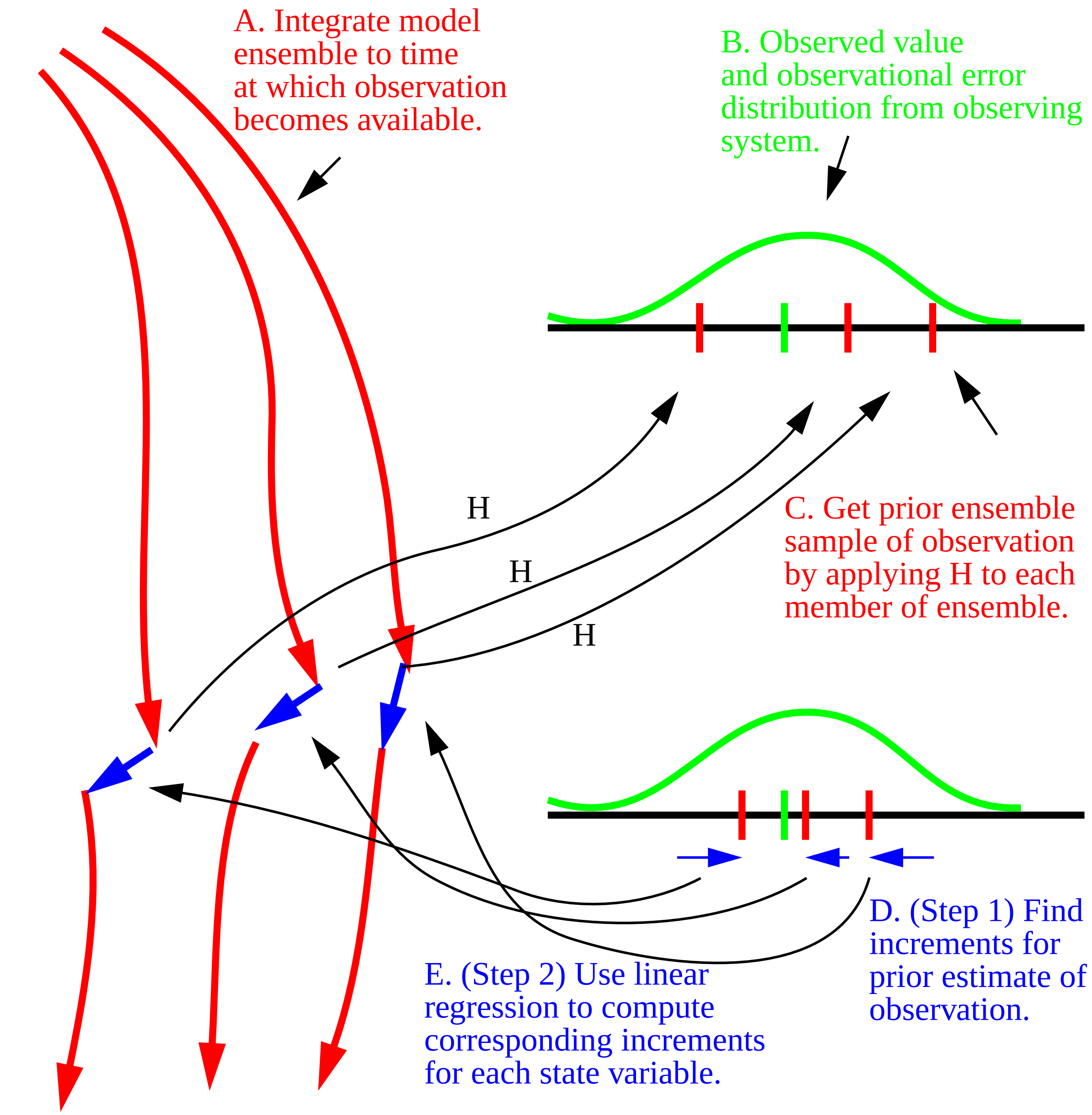
A schematic of a DART prototype.



How an Ensemble Filter Works

Theory: Impact of observations can be handled sequentially

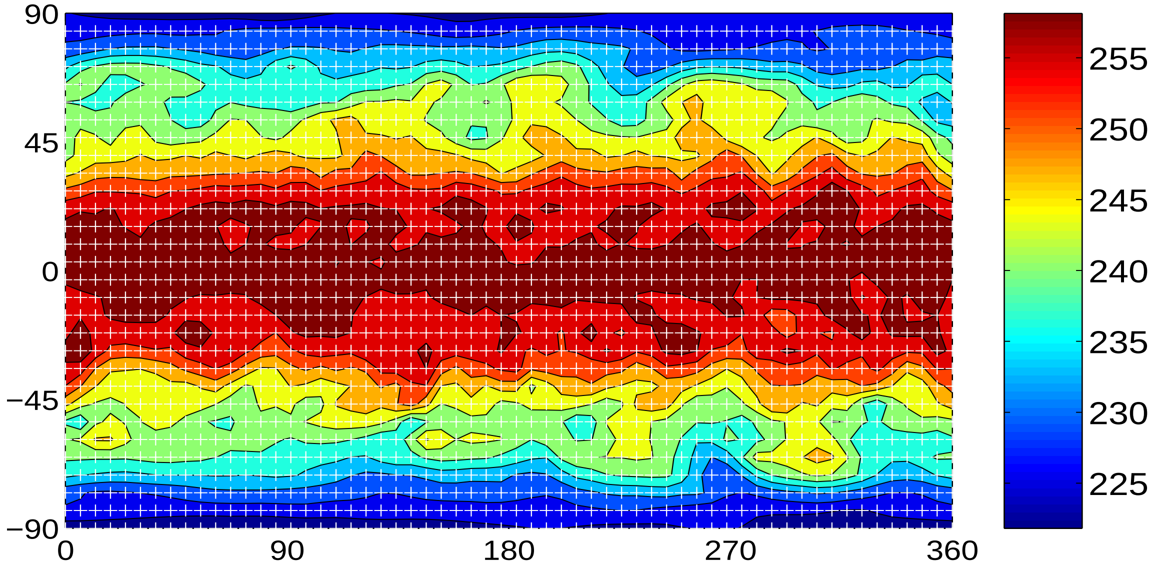
Impact of observation on each state variable can be handled sequentially



<u>GFDL FMS B-Grid Dynamical Core (Havana)</u>

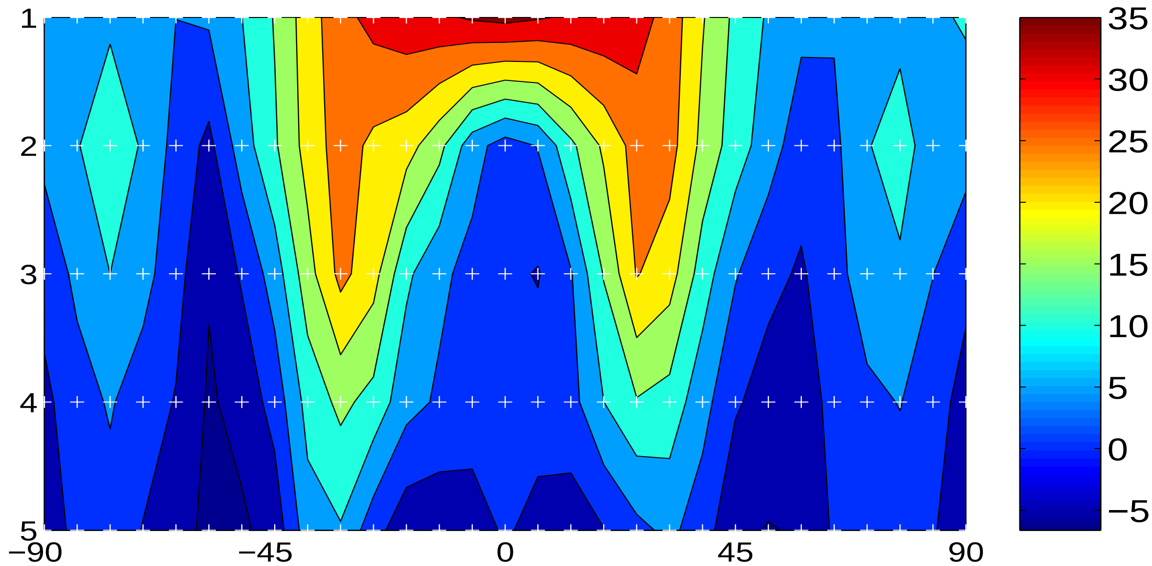
Held-Suarez Configuration (no zonal variation, fixed forcing)

T at Level 3 (Middle of Atmosphere)

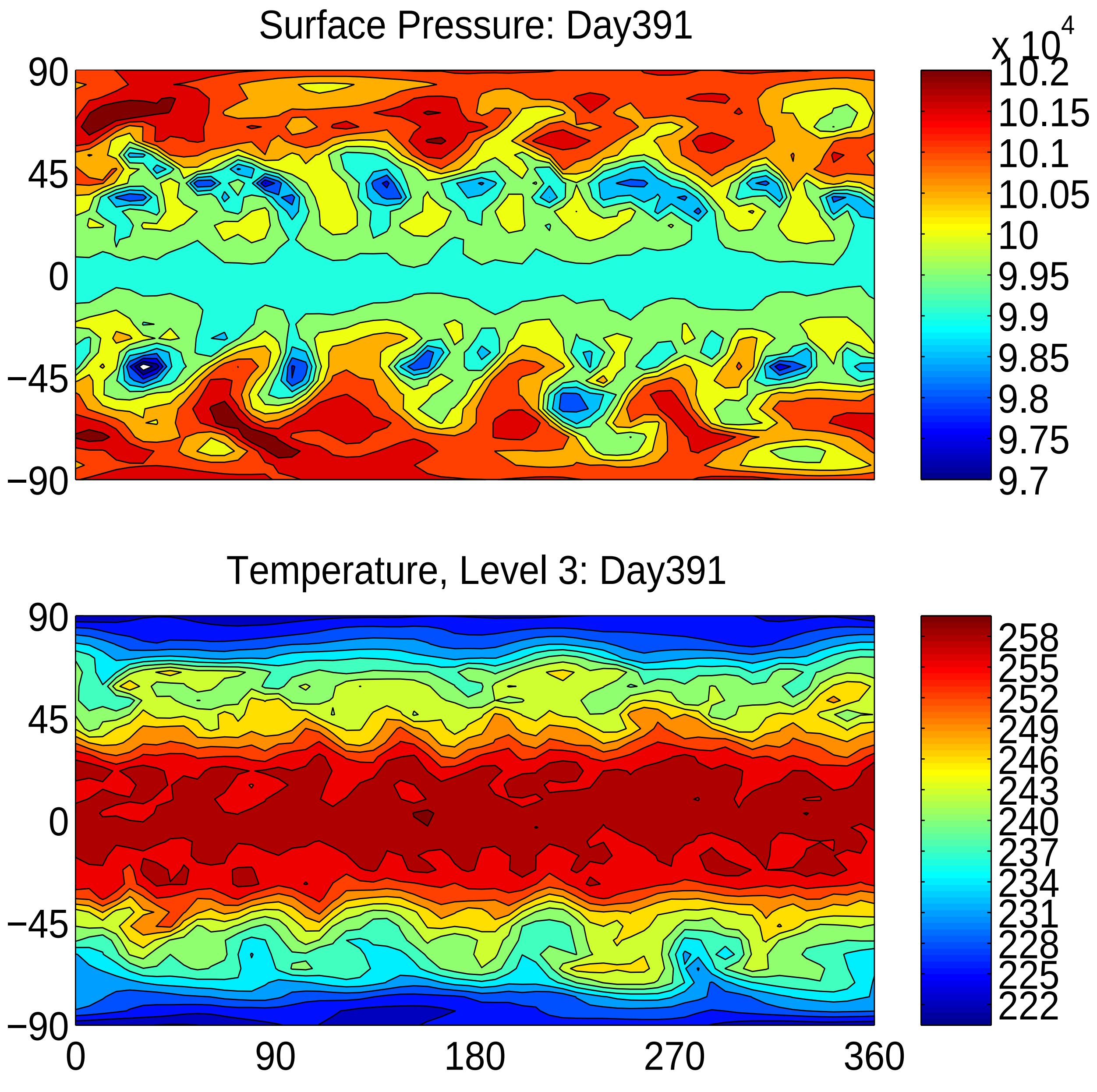


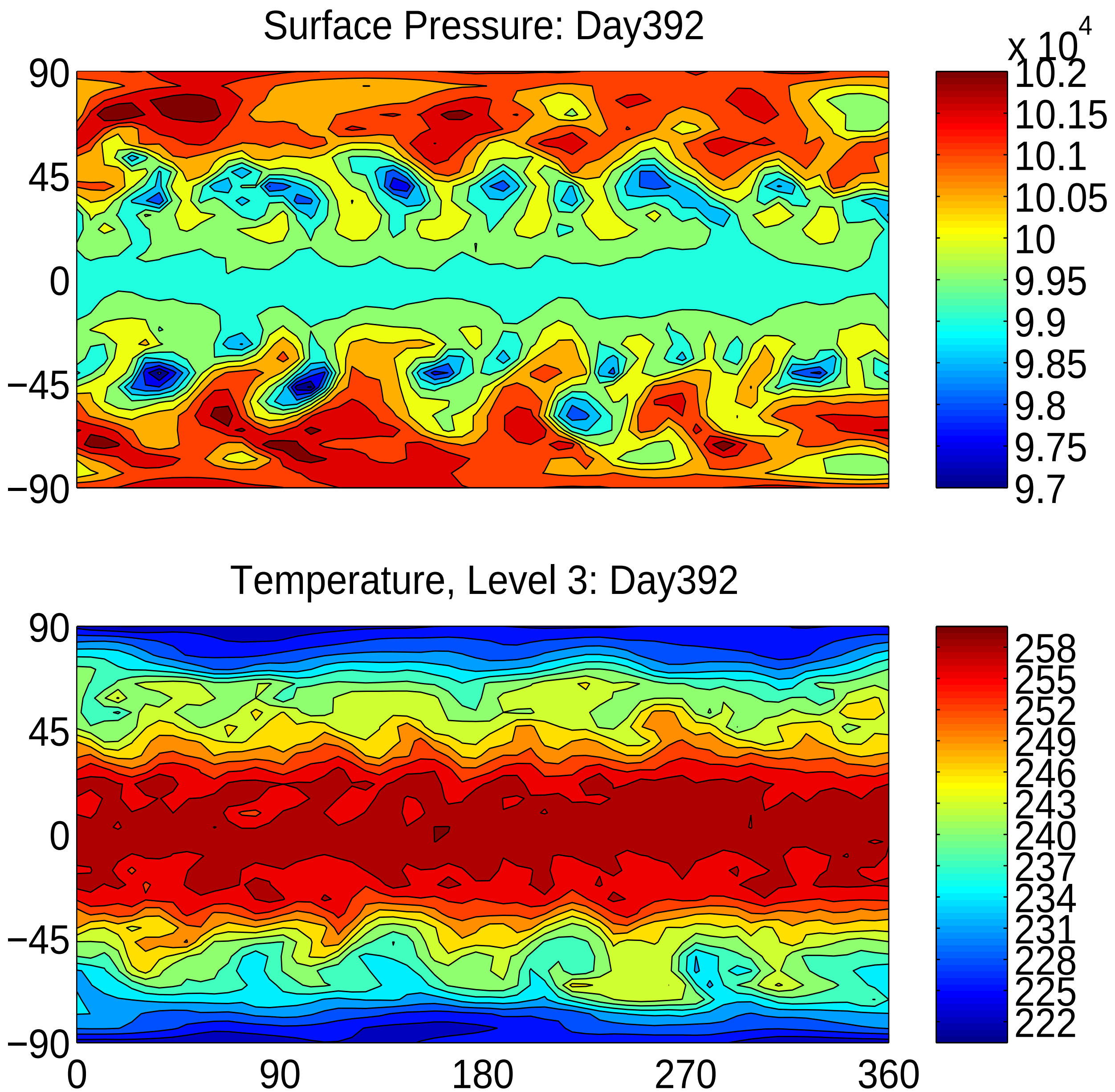
Low-Resolution (60 longitudes, 30 latitudes, 5 levels)
Damping coefficients reduced to 0.10 for error growth
Timestep 1 hour (or less for frequent observations)

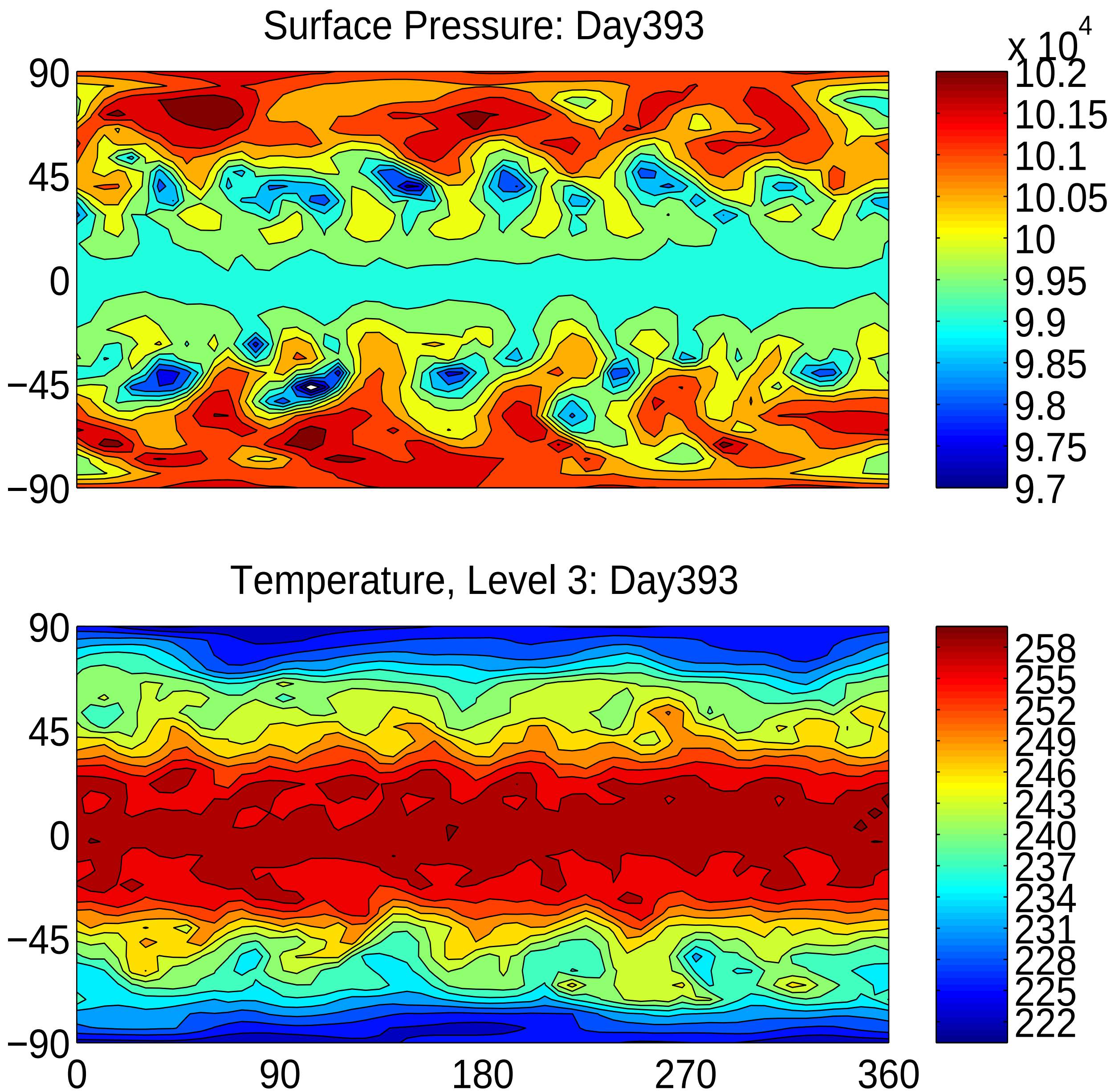
Cross section of Zonal mean U

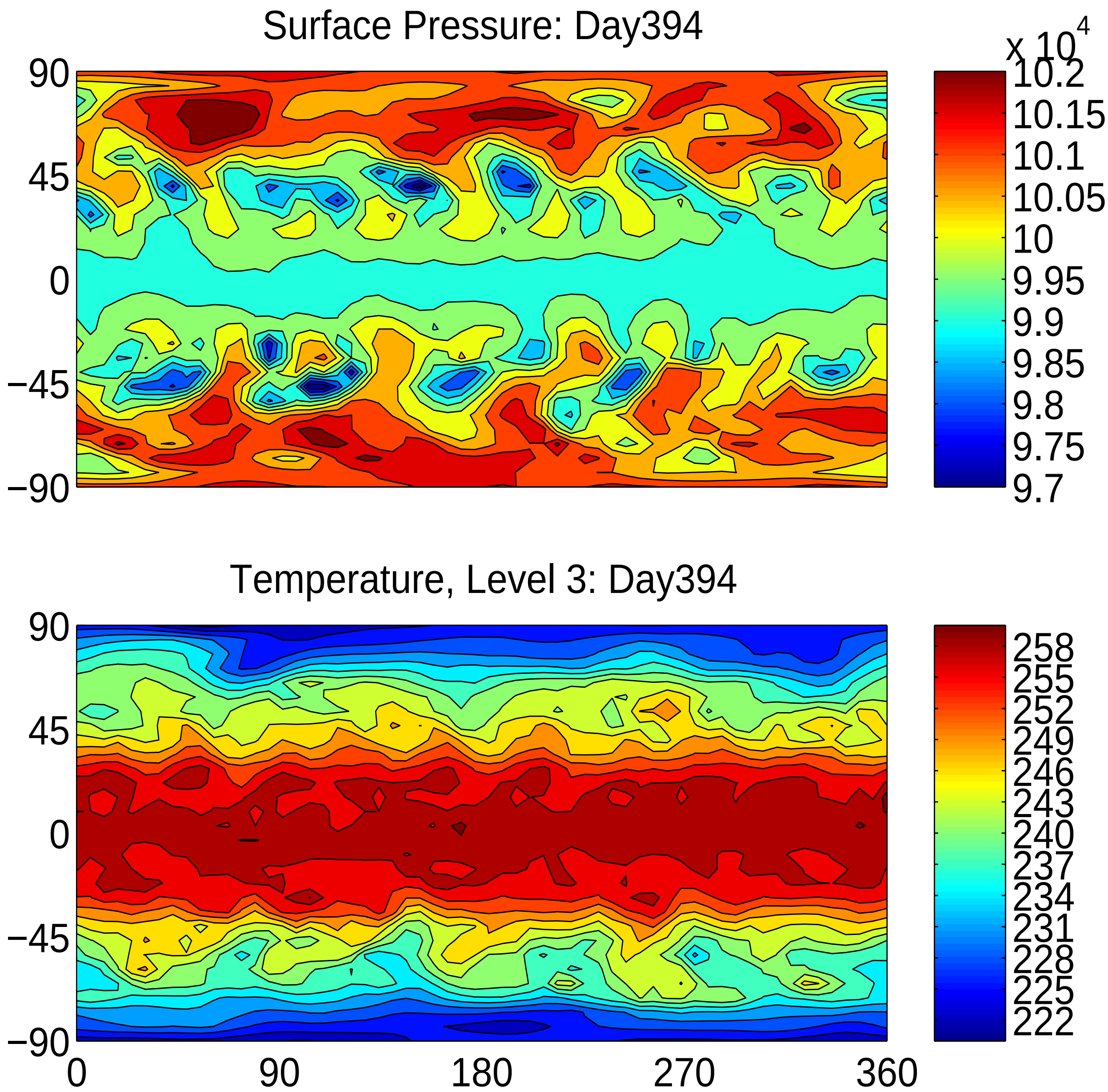


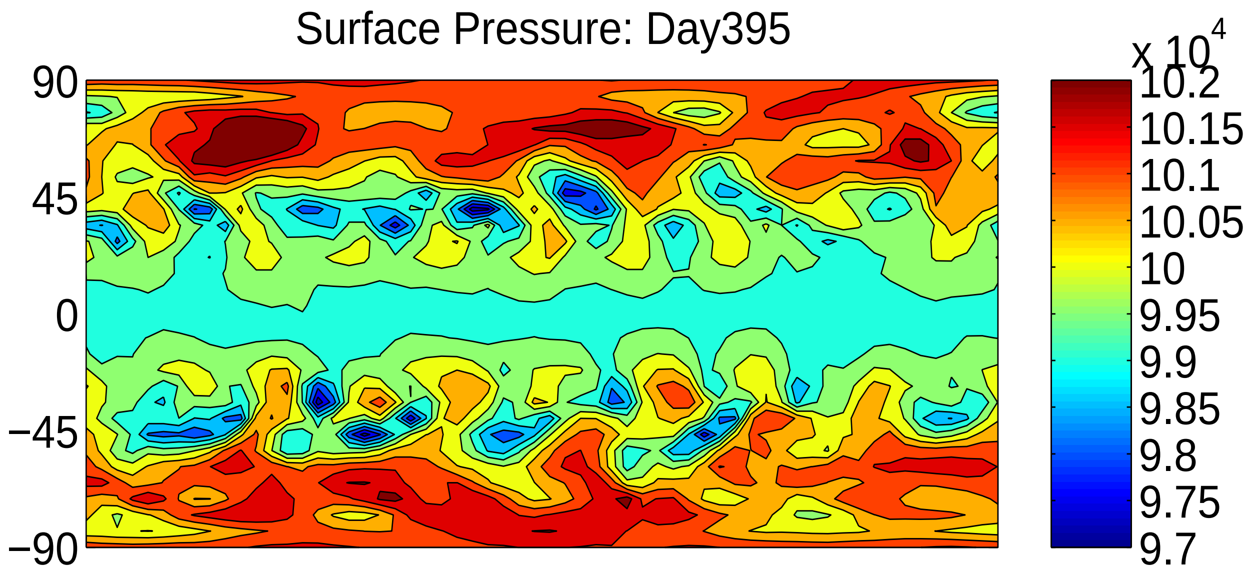
Has baroclinic instability

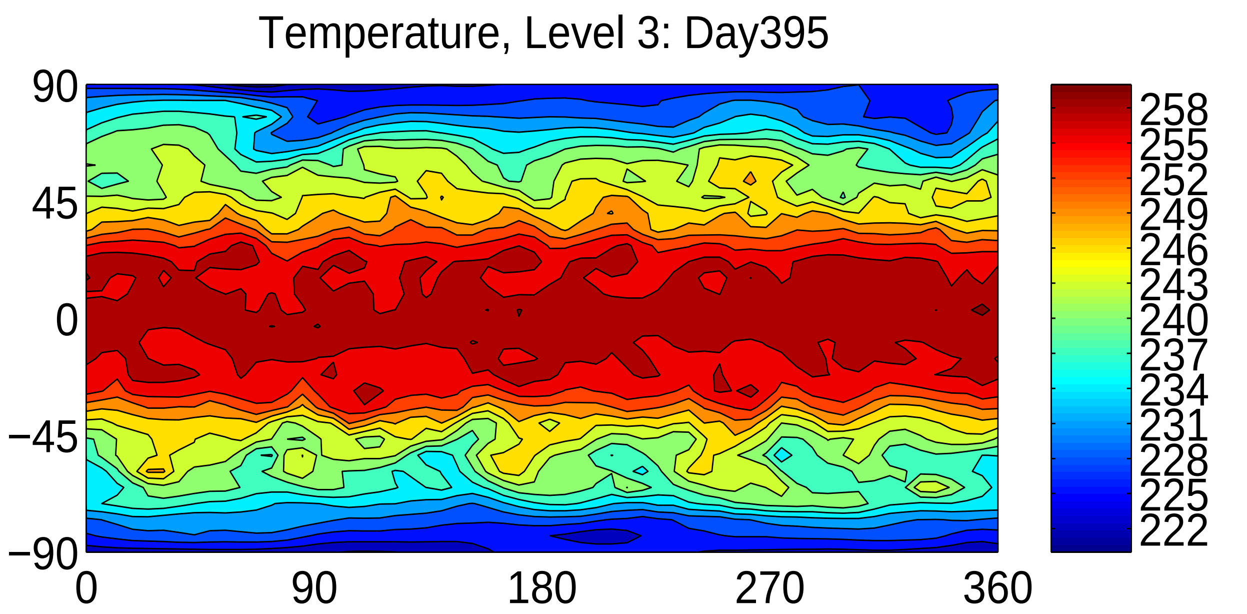

Surface Pressure: Day391
x 10^4
90
45
0
−45
−90
10.2
10.15
10.1
10.05
10
9.95
9.9
9.85
9.8
9.75
9.7
Temperature, Level 3: Day391
90
45
0
−45
−90
0
90
180
270
360
258
255
252
249
246
243
240
237
234
231
228
225
222

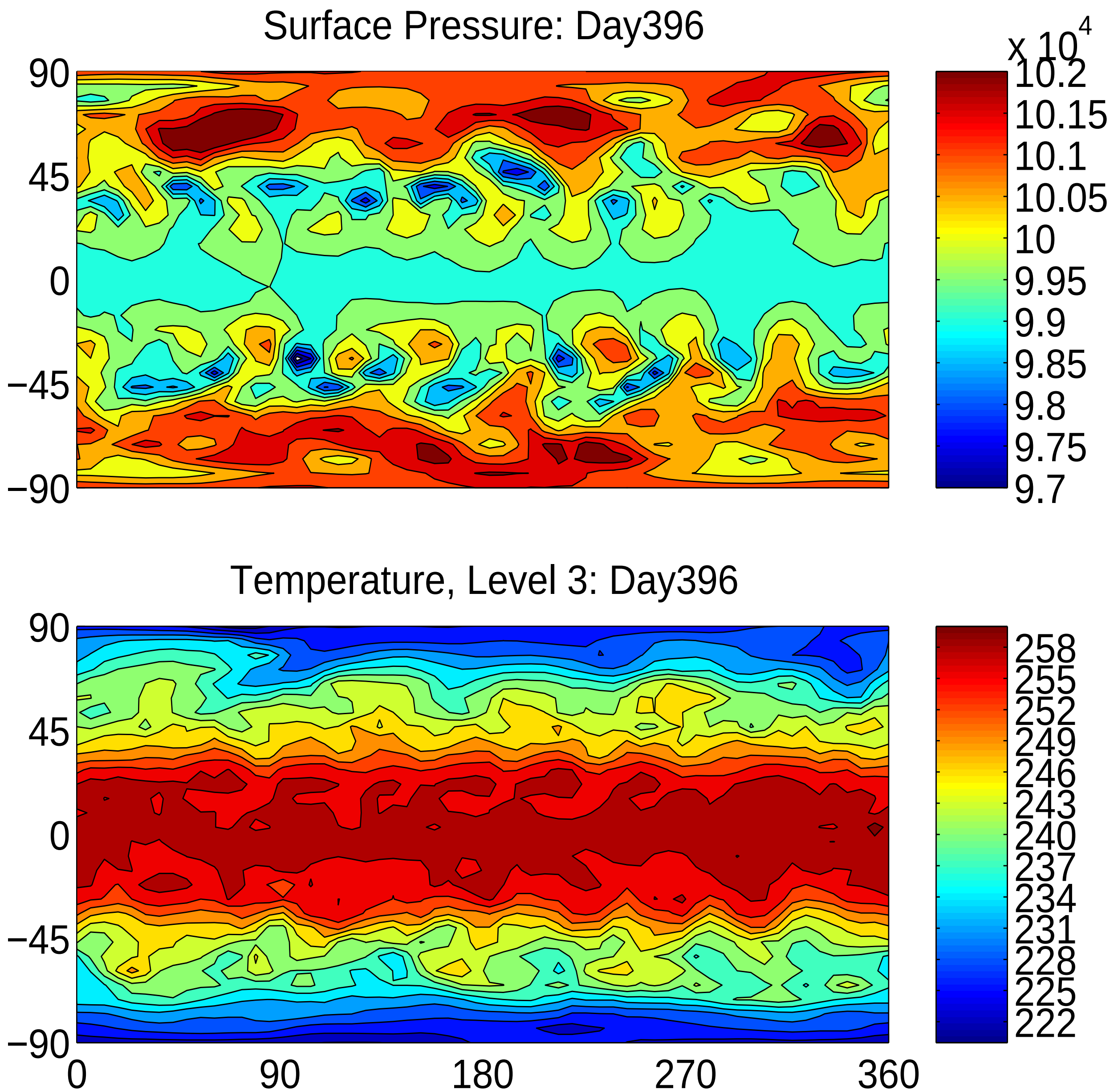

Surface Pressure: Day392
90
45
0
−45
−90
x 10^4
10.2
10.15
10.1
10.05
10
9.95
9.9
9.85
9.8
9.75
9.7
Temperature, Level 3: Day392
90
45
0
−45
−90
0
90
180
270
360
258
255
252
249
246
243
240
237
234
231
228
225
222

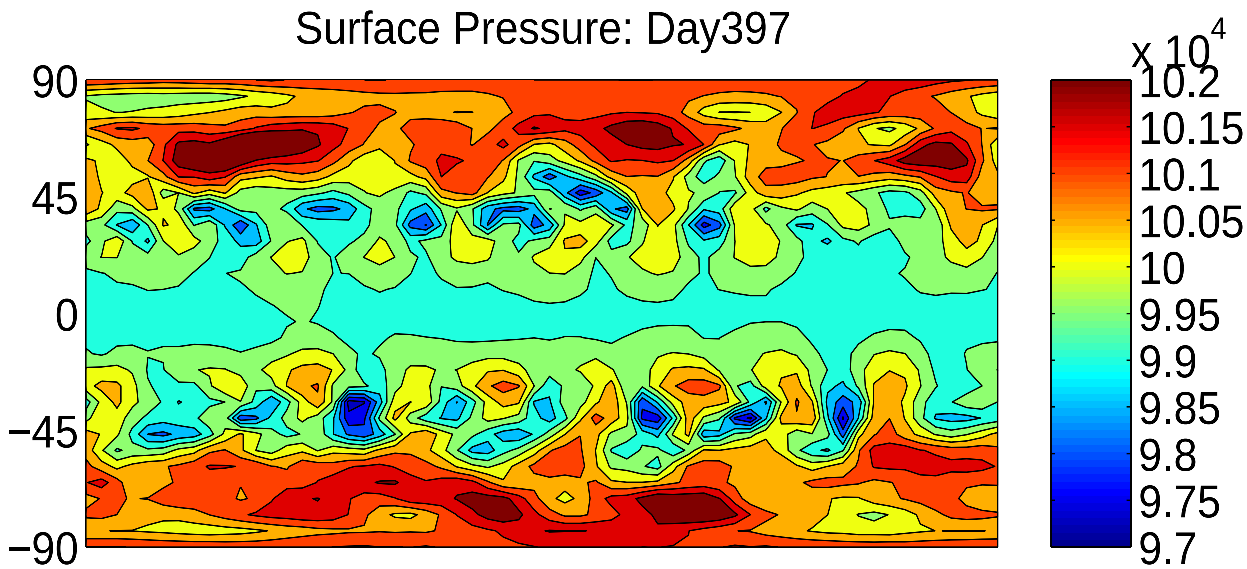

Surface Pressure: Day393
90
45
0
−45
−90
x 10^4
10.2
10.15
10.1
10.05
10
9.95
9.9
9.85
9.8
9.75
9.7
Temperature, Level 3: Day393
90
45
0
−45
−90
0
90
180
270
360
258
255
252
249
246
243
240
237
234
231
228
225
222

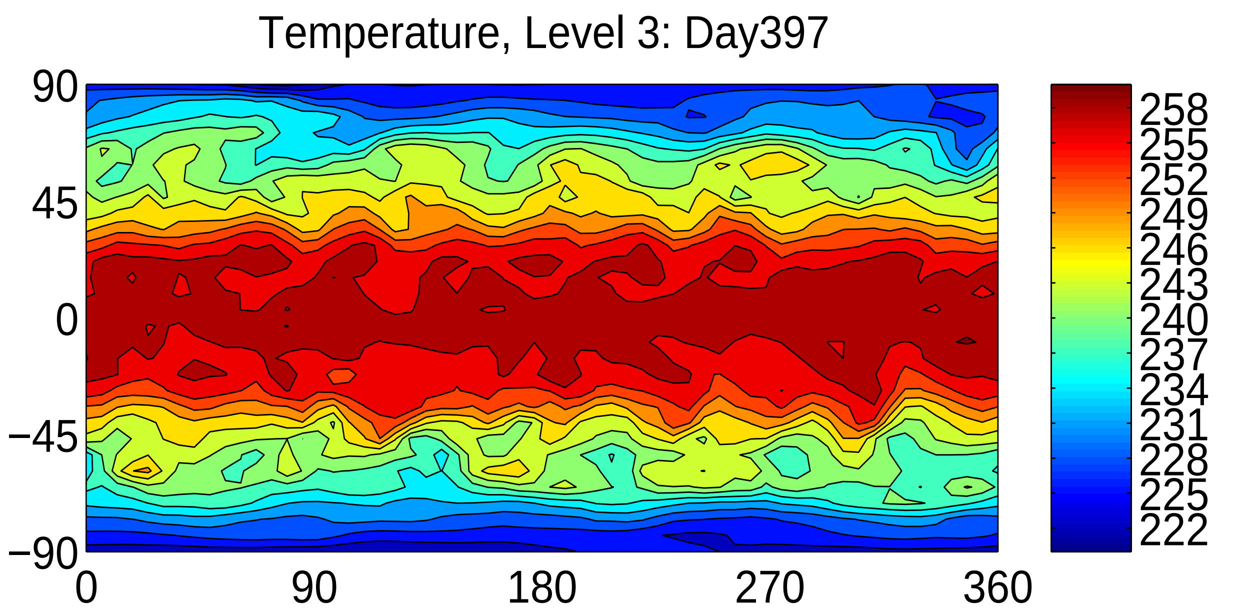

Surface Pressure: Day394
x 10^4
10.2
10.15
10.1
10.05
10
9.95
9.9
9.85
9.8
9.75
9.7
90
45
0
−45
−90
Temperature, Level 3: Day394
258
255
252
249
246
243
240
237
234
231
228
225
222
90
45
0
−45
−90
0
90
180
270
360

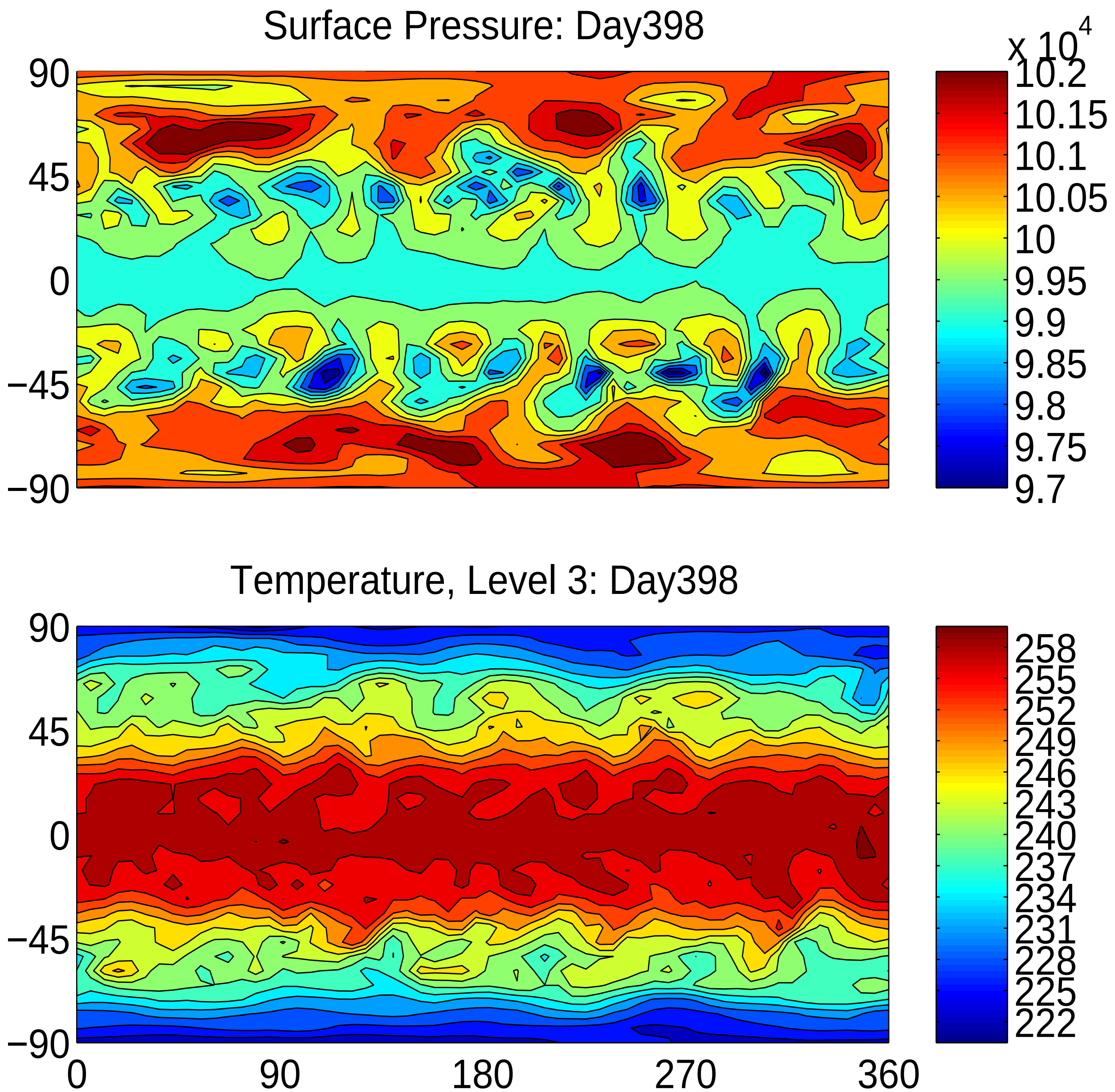

Surface Pressure: Day395
90
45
0
−45
−90
x 10^4
10.2
10.15
10.1
10.05
10
9.95
9.9
9.85
9.8
9.75
9.7

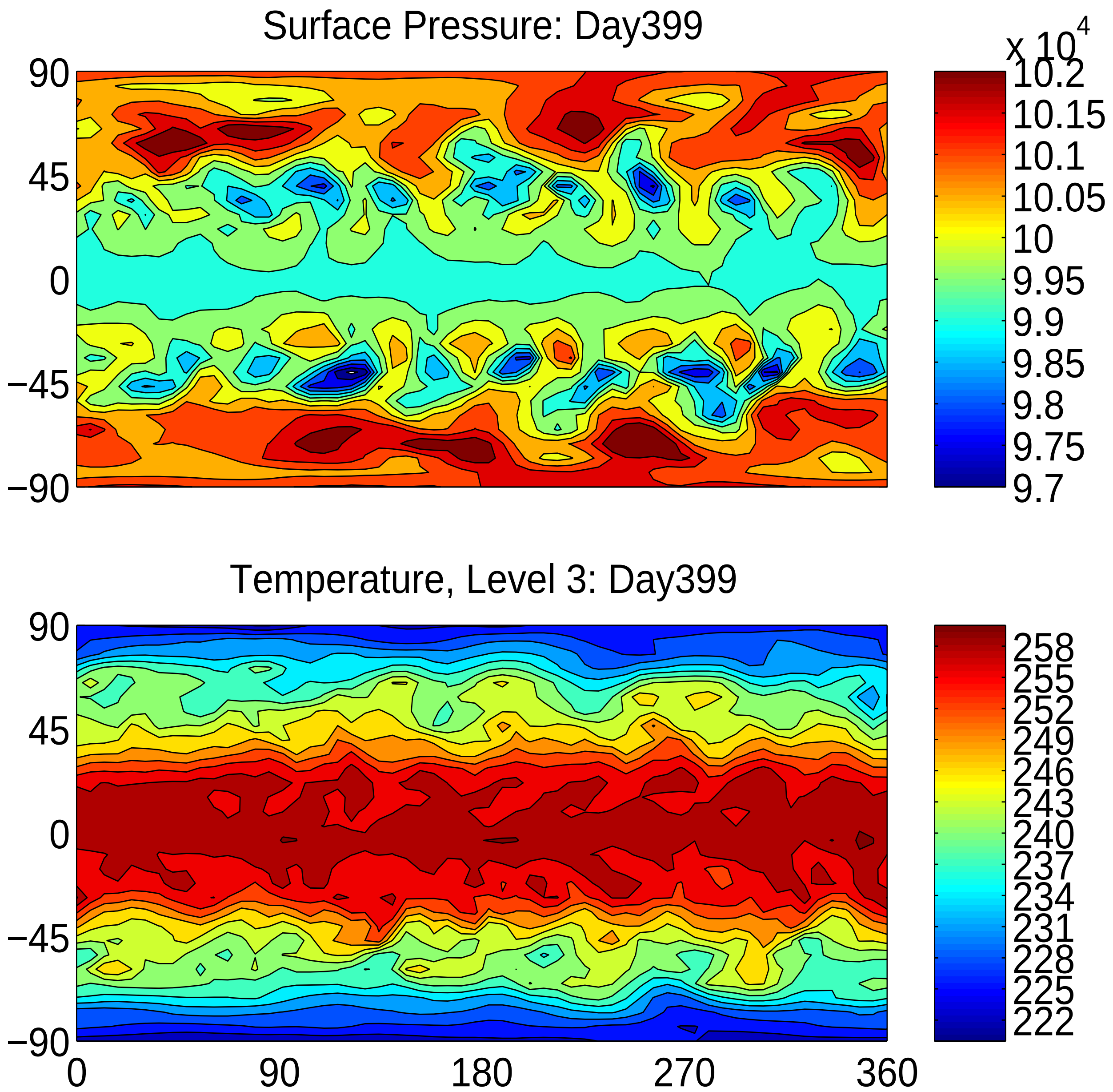

Temperature, Level 3: Day395
90
45
0
−45
−90
0
90
180
270
360
258
255
252
249
246
243
240
237
234
231
228
225
222

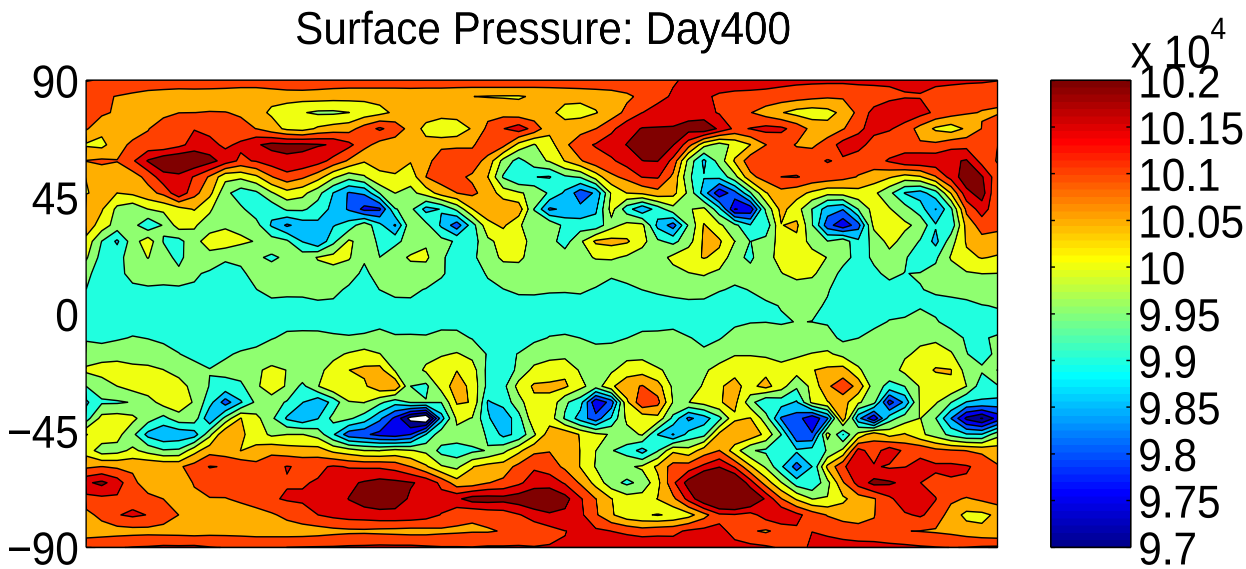

Surface Pressure: Day396
x 10^4
10.2
10.15
10.1
10.05
10
9.95
9.9
9.85
9.8
9.75
9.7
90
45
0
−45
−90
Temperature, Level 3: Day396
258
255
252
249
246
243
240
237
234
231
228
225
222
90
45
0
−45
−90
0
90
180
270
360

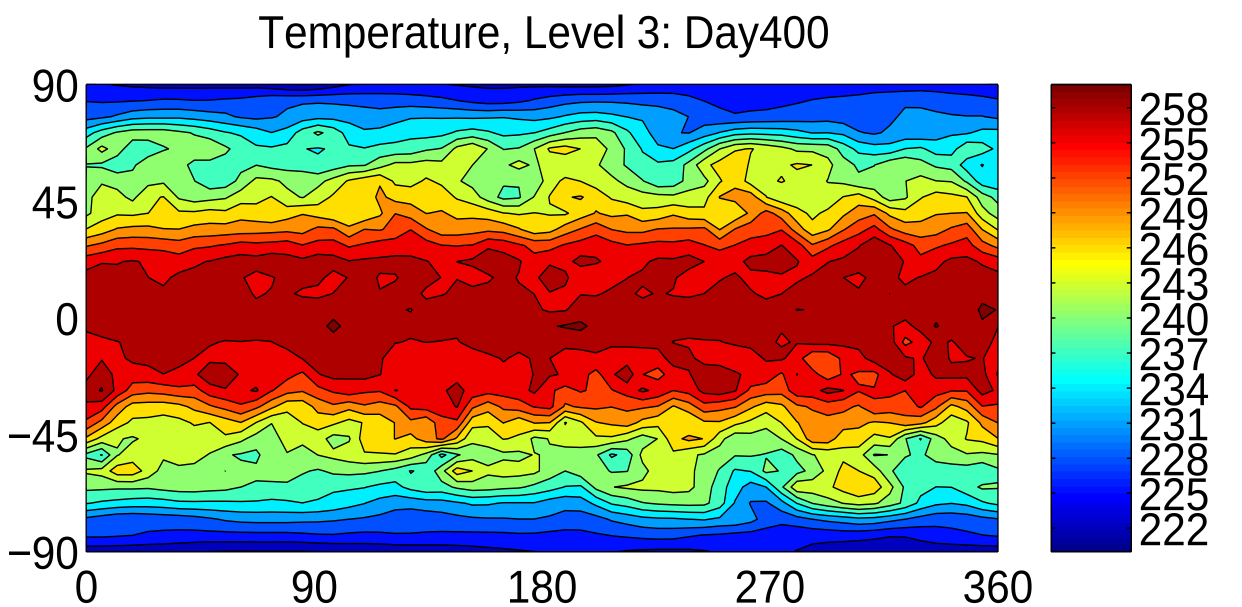

Surface Pressure: Day397
x 10^4
10.2
10.15
10.1
10.05
10
9.95
9.9
9.85
9.8
9.75
9.7
90
45
0
−45
−90


Temperature, Level 3: Day397
258
255
252
249
246
243
240
237
234
231
228
225
222
90
45
0
−45
−90
0
90
180
270
360


Surface Pressure: Day398
x 10^4
10.2
10.15
10.1
10.05
10
9.95
9.9
9.85
9.8
9.75
9.7
90
45
0
−45
−90
Temperature, Level 3: Day398
258
255
252
249
246
243
240
237
234
231
228
225
222
90
45
0
−45
−90
0
90
180
270
360


Surface Pressure: Day399
90
45
0
−45
−90
x 10^4
10.2
10.15
10.1
10.05
10
9.95
9.9
9.85
9.8
9.75
9.7
Temperature, Level 3: Day399
90
45
0
−45
−90
0
90
180
270
360
258
255
252
249
246
243
240
237
234
231
228
225
222

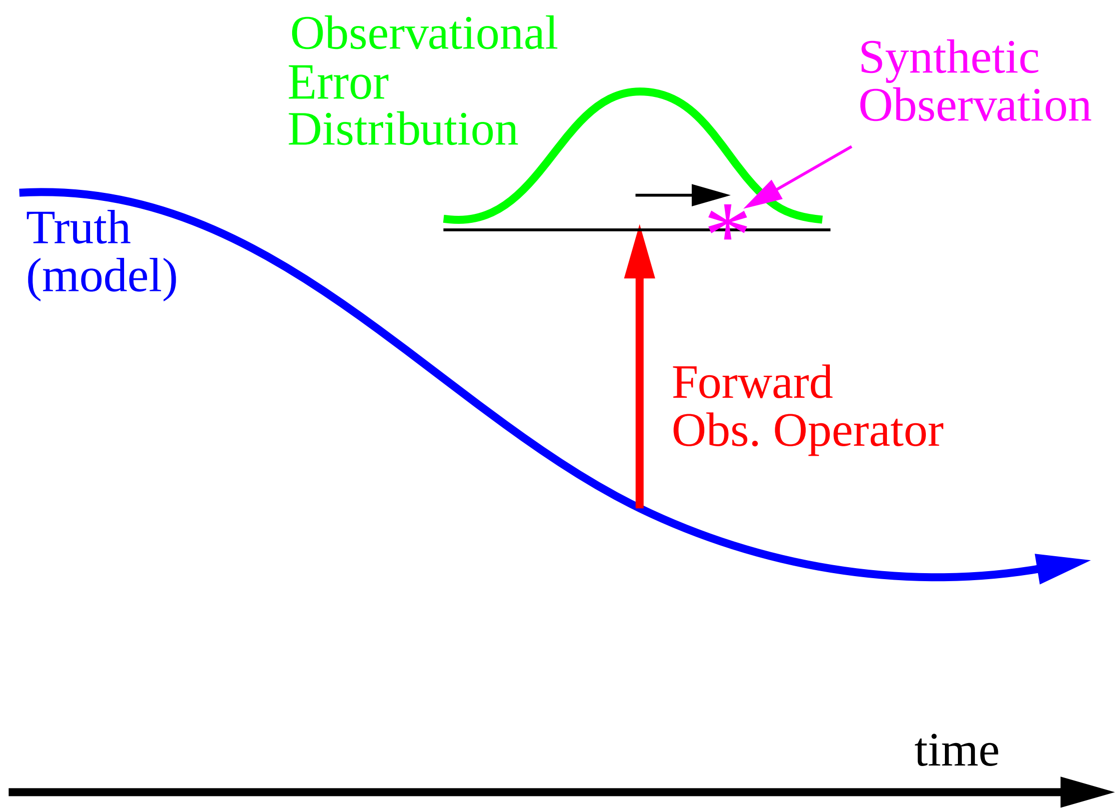

Surface Pressure: Day400
90
45
0
−45
−90
x 10^4
10.2
10.15
10.1
10.05
10
9.95
9.9
9.85
9.8
9.75
9.7


Temperature, Level 3: Day400
90
45
0
−45
−90
0
90
180
270
360
258
255
252
249
246
243
240
237
234
231
228
225
222

Perfect model experiments

'Truth' is generated by integrating model

B-grid, integrated for 100 years from state of rest before starting
(Multi-year spin-up for upper level temperatures)



'Synthetic' obs. by applying 'forward observation' operator to truth
(Here, this is just interpolating to a random horizontal location)

Instrument error simulated by adding random draw from a specified Gaussian distribution to the interpolated observation

All the assimilation algorithm ever sees is these simulated observations

Result of assimilation can be compared to 'truth'

Experimental Design Details

Base case assimilation starts from ‘climatological’ ensemble

Add tiny perturbations to control integration (truth)
Integrate this ensemble for several years

Ensemble size is 20 for ALL cases here

Each assimilation case is run for 400 days

Summary results are from last 200 days

No bias correction steps taken (no covariance inflation)

Single tuning parameter controls distance dependent correlation mask
Gives less weight to distant observations
This was tuned to give best RMS results in base case
Not changed for any other experiments

Note: Level 1 temperature in Held-Suarez configuration has very low frequency adjustment,

Experimental Sets

1. Impact of spatial density of observations:
   150, 300, 450, 900, 1800, 3600, 7200, 14400, 28800 PS obs
   Every 24 hours
   PS observational error standard deviation 1.0 mb

2. Impact of frequency of observations
   1800 PS observations
   Every 24, 12, 6, 4, 3, 2, and 1 hours, 30, 15, and 5 minutes
   PS observational error standard deviation 1.0 mb

3. Information content of different observation types
   1800 observations of PS, or low-level T, or low-level U/V
   Every 24 hours
   PS observational error SD 2.0 and 1.0 mb
   T observational error SD 1.0 and 0.5 K
   U/V observational error SD 2.0 and 1.0 m/s, U, V errors independent

4. What happens if observations are confined to limited spatial domain
   450 PS obs, only in N. Hemisphere between 90 and 270 deg. longitude
   Every 24 hours
   PS observational error standard deviation 1.0 mb

5. Impact of increased vertical resolution
   1800 PS obs
   Every 24 hours
   PS observational error standard deviation 1.0 mb
   5 and 18 vertical levels

6. Impact of adding stochastic ‘sub-grid scale’ noise
   1800 PS obs, Every 24 hours
   PS observational error standard deviation 1.0
   Temperature time tendency noise standard deviation 0, 10%, 40%

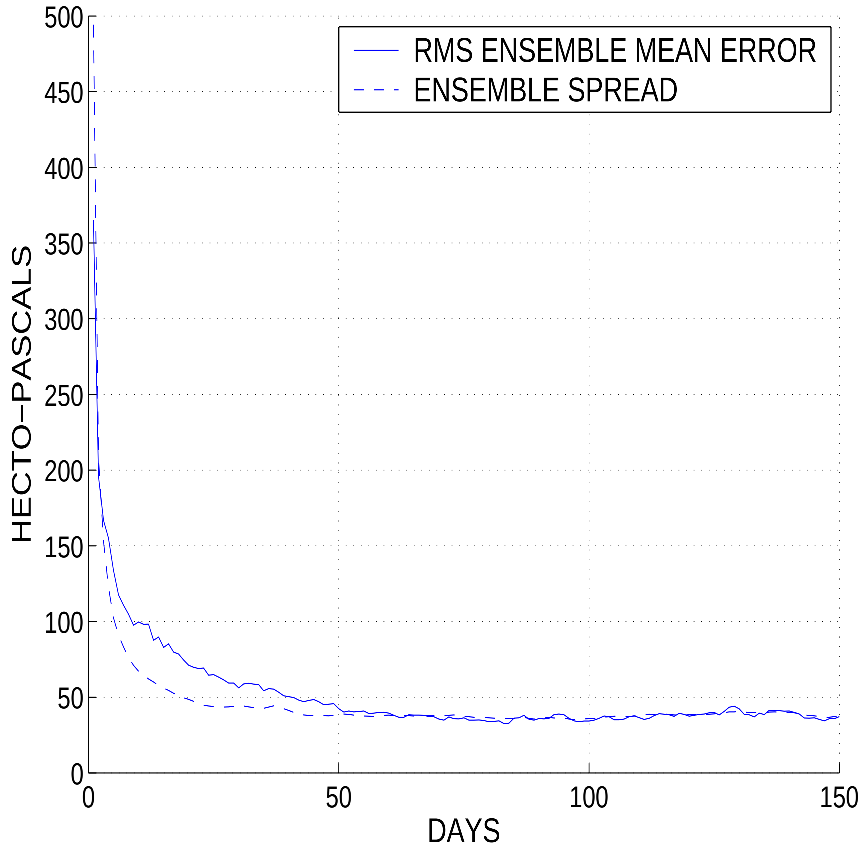
Baseline Case: 1800 PS Obs every 24 hours

PS Error Reduces by about factor of 10

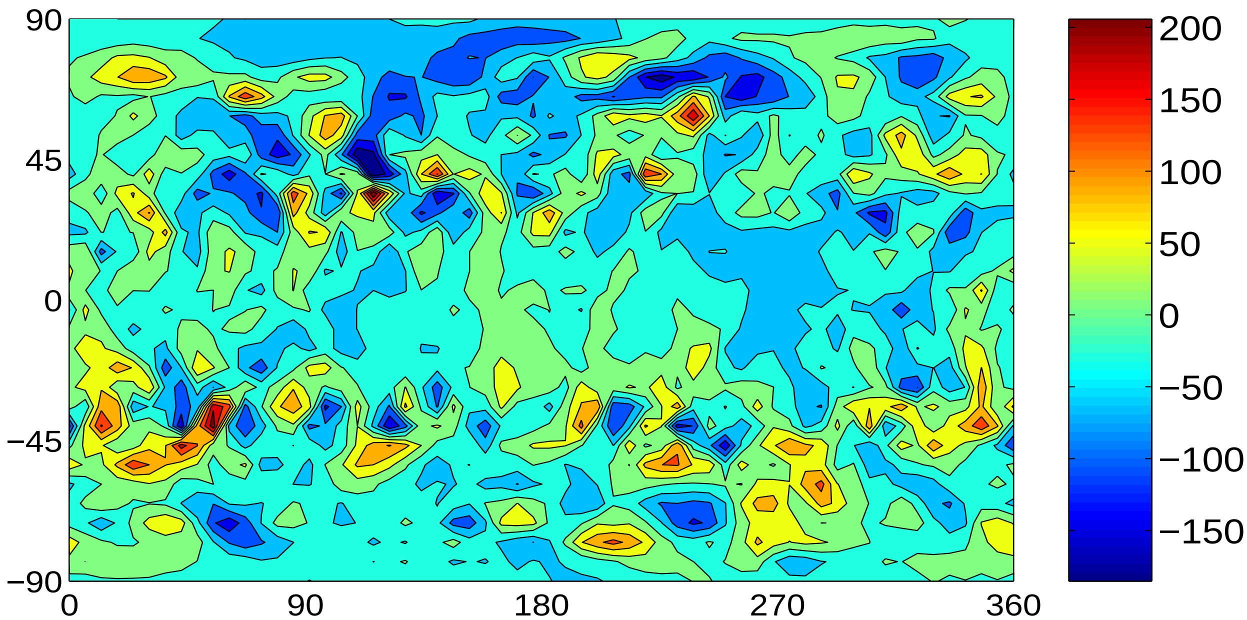
Asymptotes after about 50 days

Ensemble spread is approximately correlated with RMS error

Final Error is about 0.4 mb



Largest error in mid-latitudes, ‘synoptic’ scales after 400 days

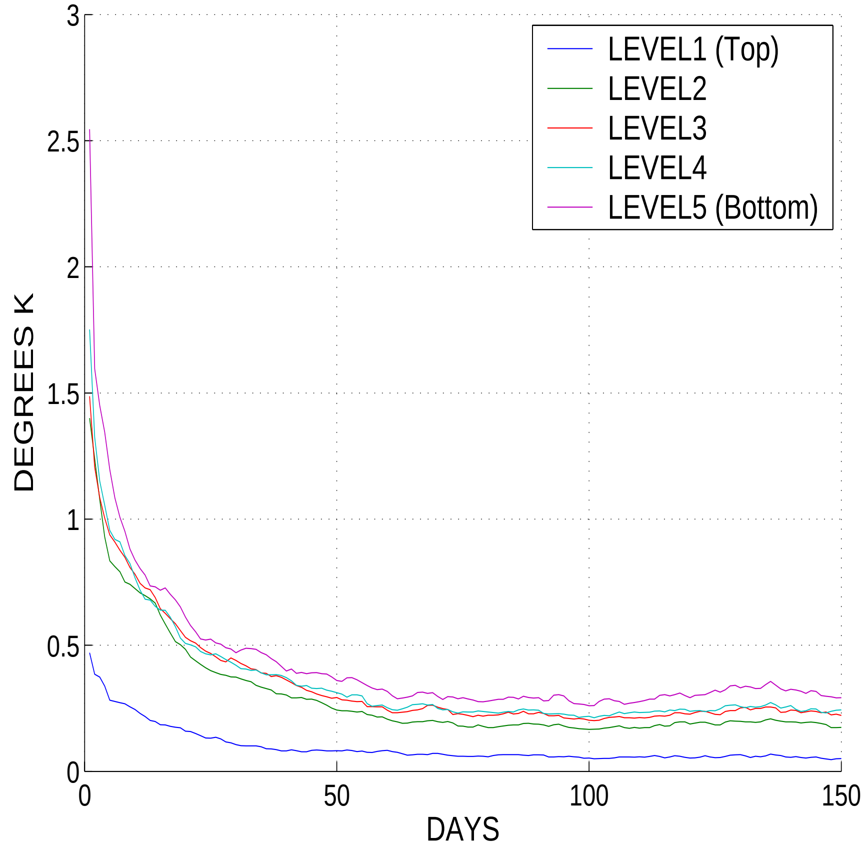


Baseline Case: 1800 PS Obs every 24 hours

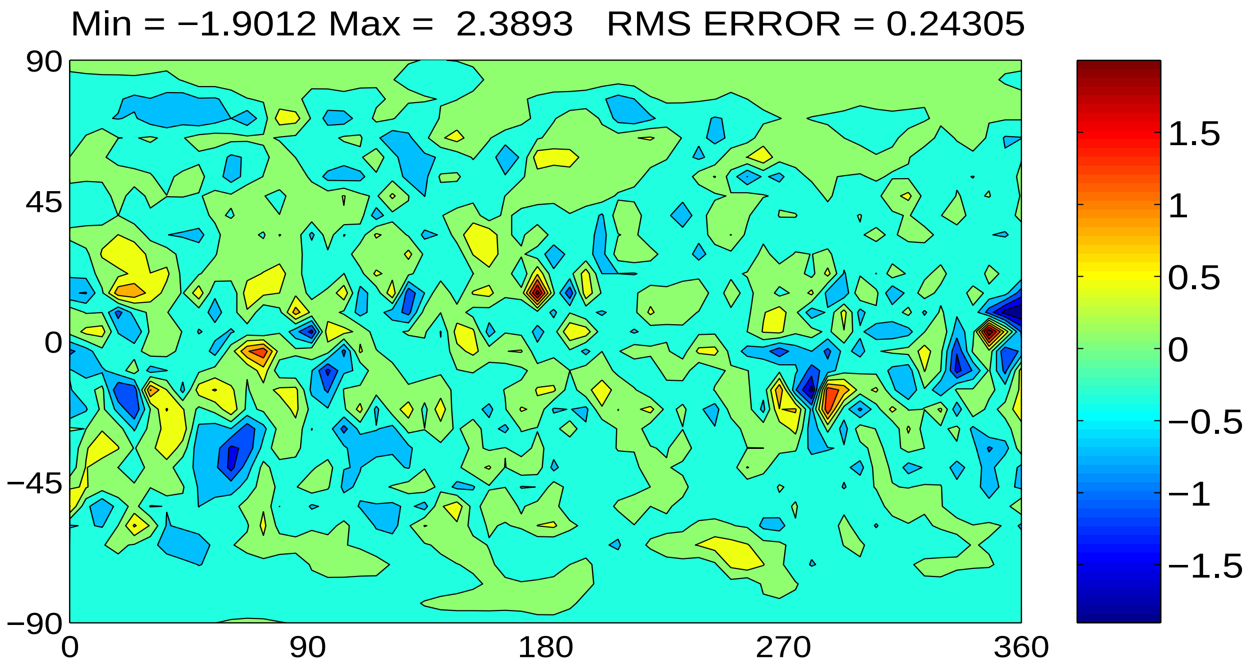
T Error also reduced by about factor of 10

Asymptotes about day 70

Final error about 0.25 K for interior levels

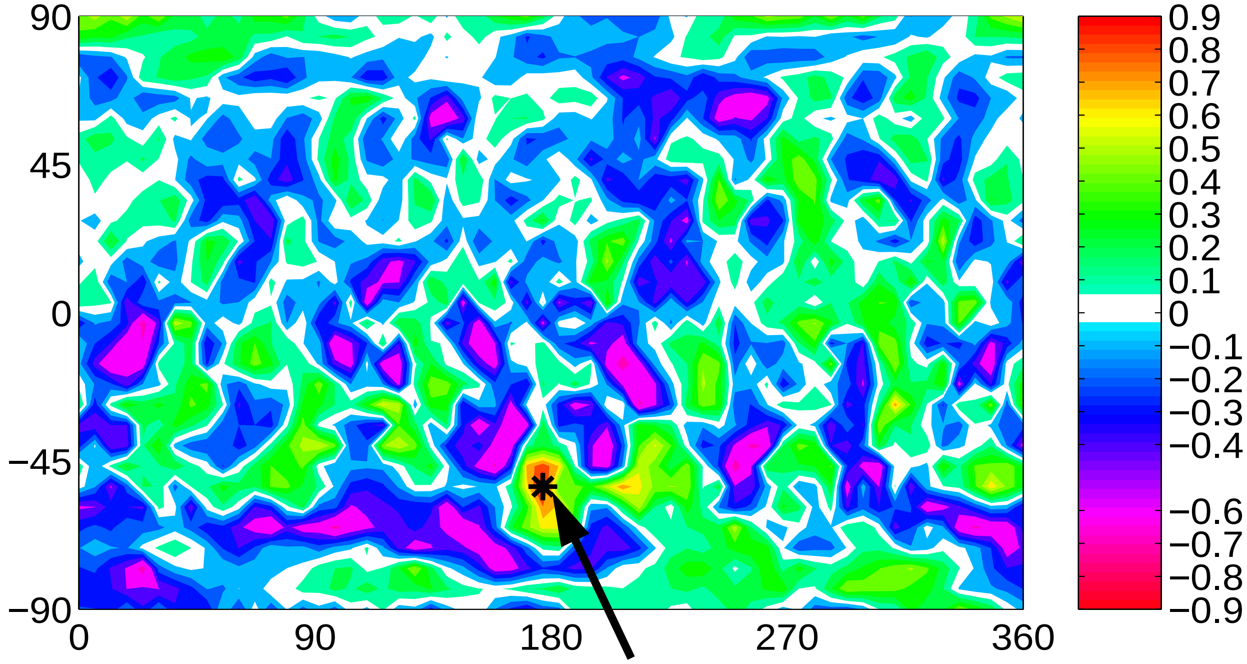


Largest T error in tropics for interior levels (level 3, day 400 shown)

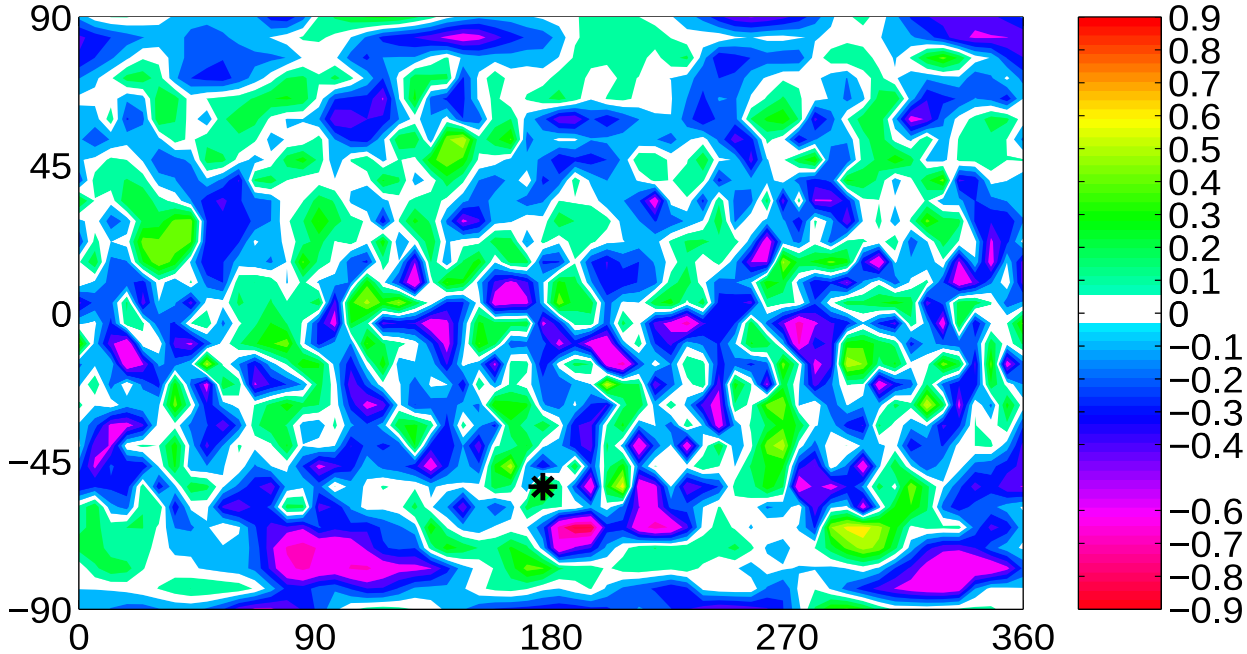


Sample Correlation: Baseline Case

Sample correlations reflect how observations can impact state variables:



Correlation of PS with PS at (180, 50S): largest values local but noisy



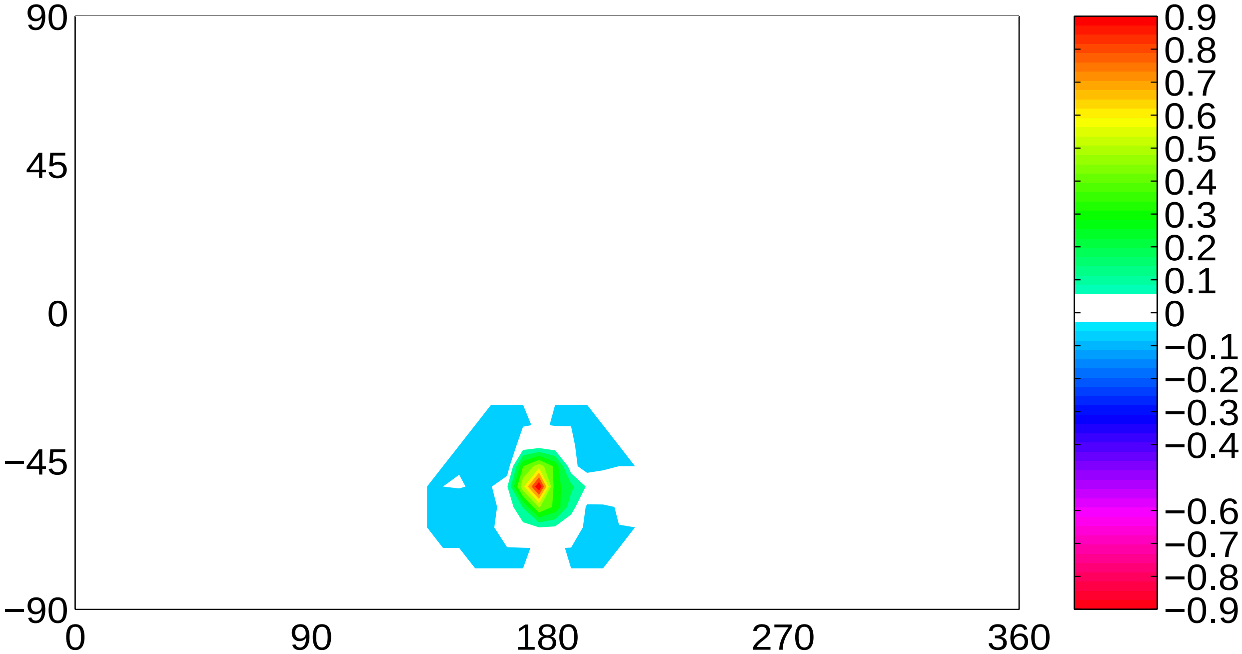
Correlation of T at level 3 with PS at (180, 50S);

Lots of noise, limited local signal

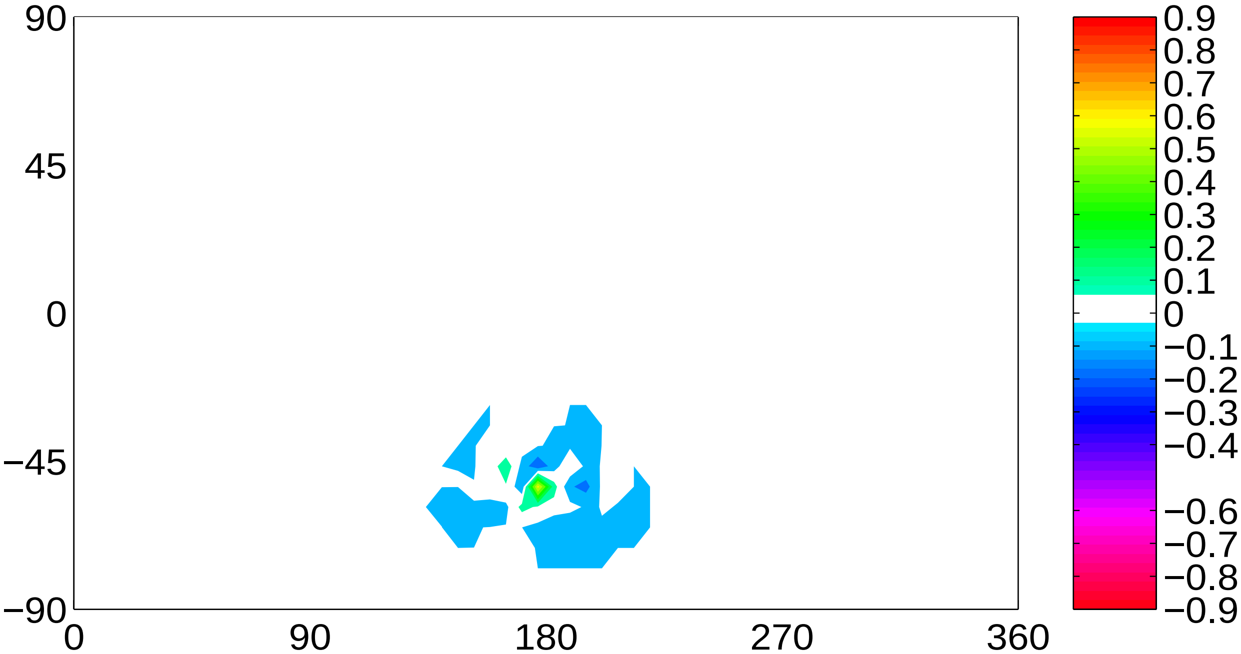
Filter must be able to extract limited signal from lots of noise

Sample Correlation with Envelope: Baseline Case

Sample correlations reflect how observations can impact state variables:

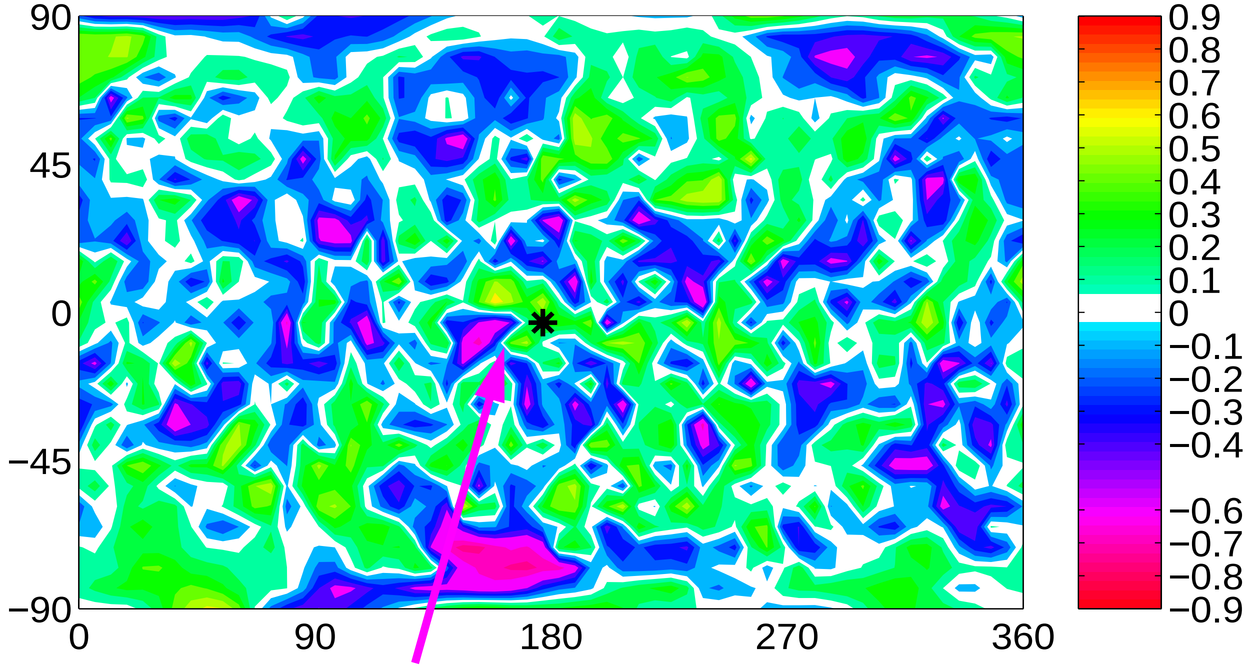


Correlation of PS with PS at (180, 50S):

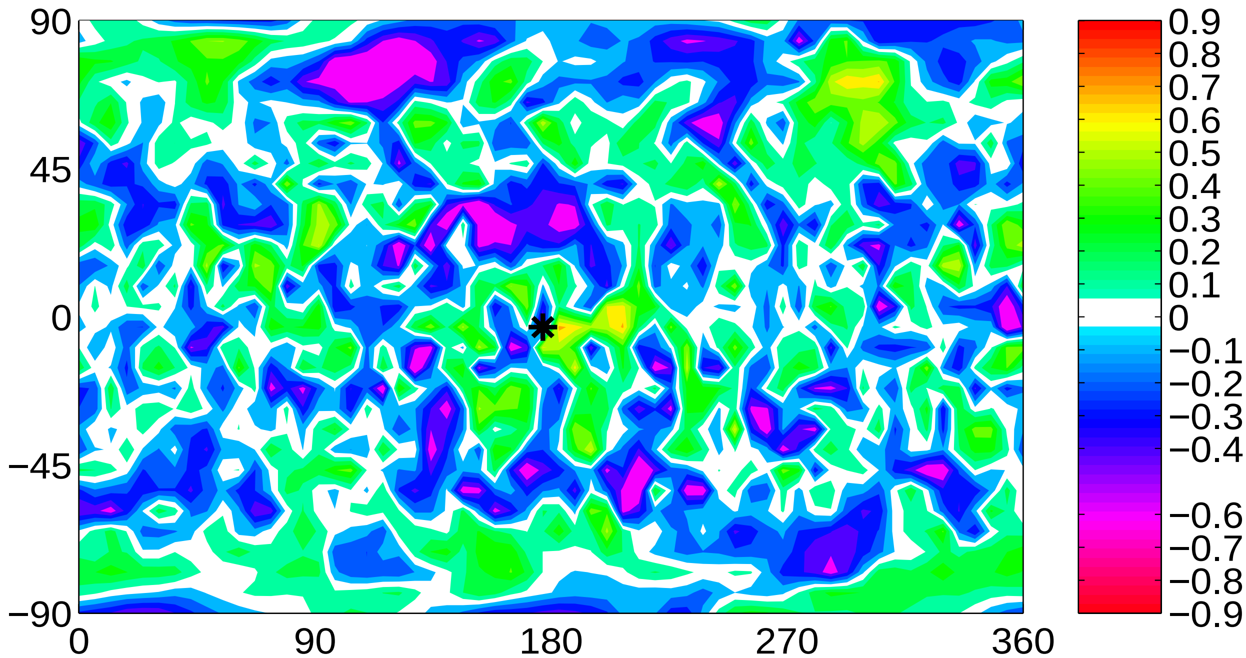


Correlation of T at level 3 with PS at (180, 50S);

Sample correlations vary significantly in time and space

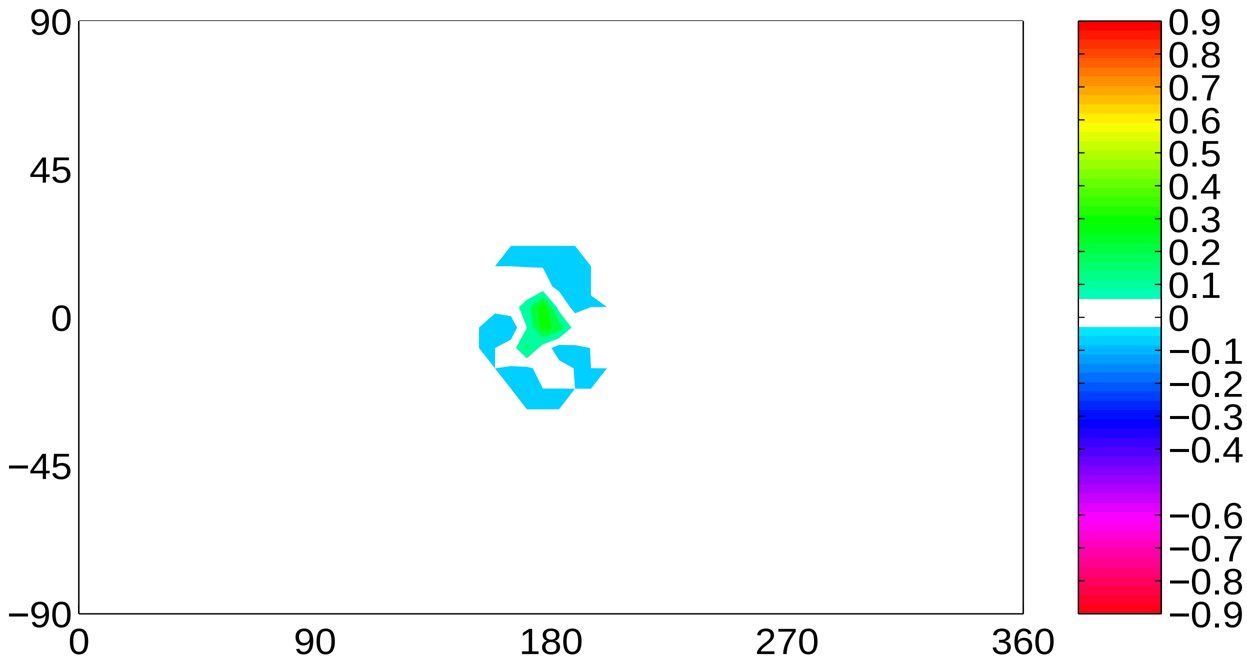


Correlation of PS at (180, 0) with T at level 3

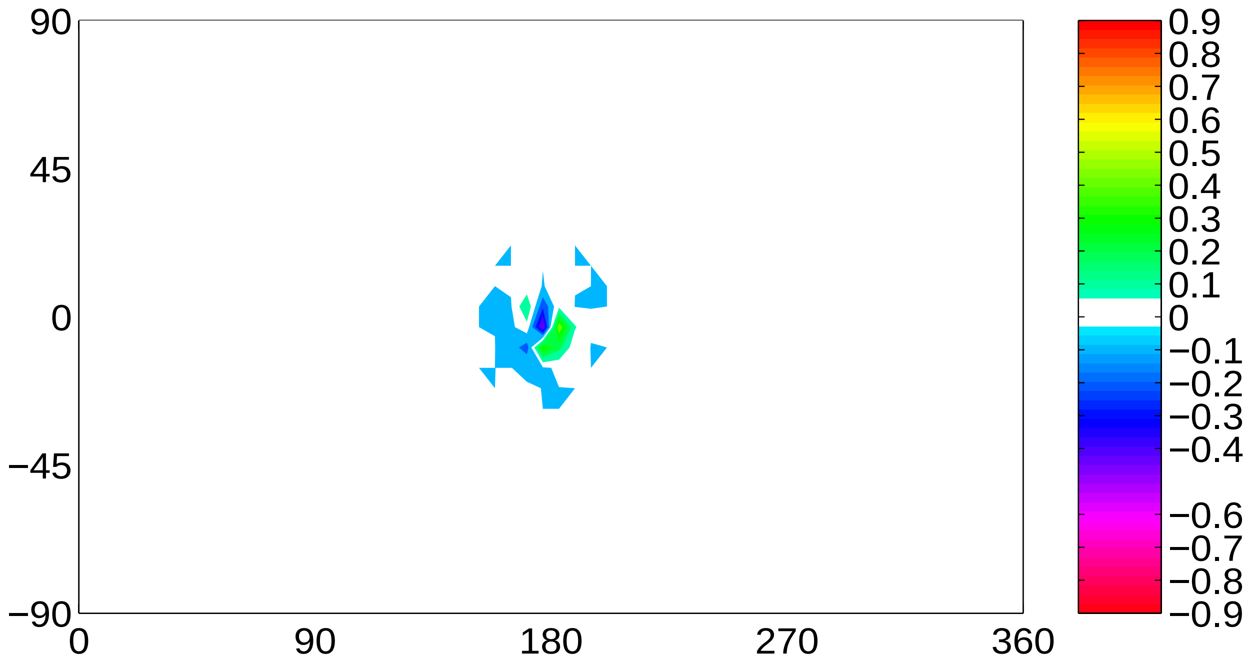


Same field, but 10 days later; Local structure is somewhat similar
Noise at a distance has moved around randomly
<u>Must take actions to avoid impact from remote noise</u>

Sample correlations vary significantly in time and space



Correlation of PS at (180, 0) with T at level 3



Same field, but 10 days later

Impacts of spatial density of PS obs

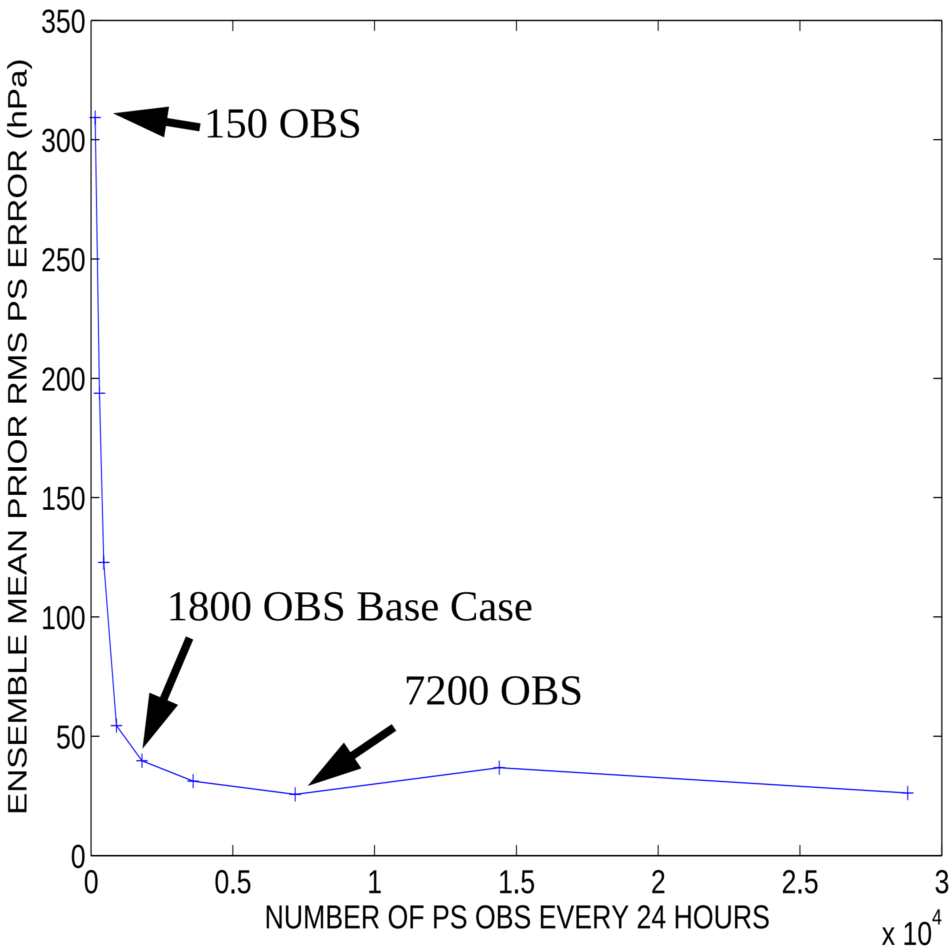
150, 300, 450, 900, 1800, 3600, 7200, 14400 and 28,800 every 24 hours

PS error reduces to about 0.3 mb

150 obs reduces climatological error by less than half

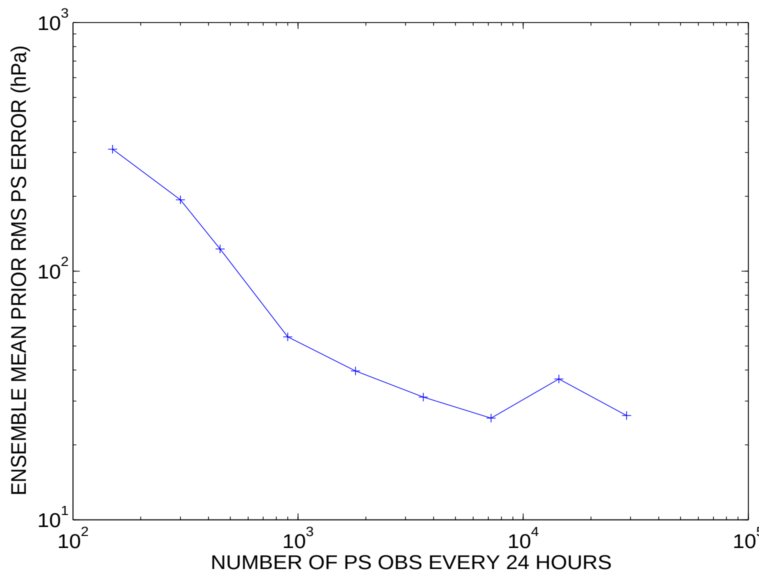
More than 7200 obs appears superfluous

Why is 14,400 worse? No clue.



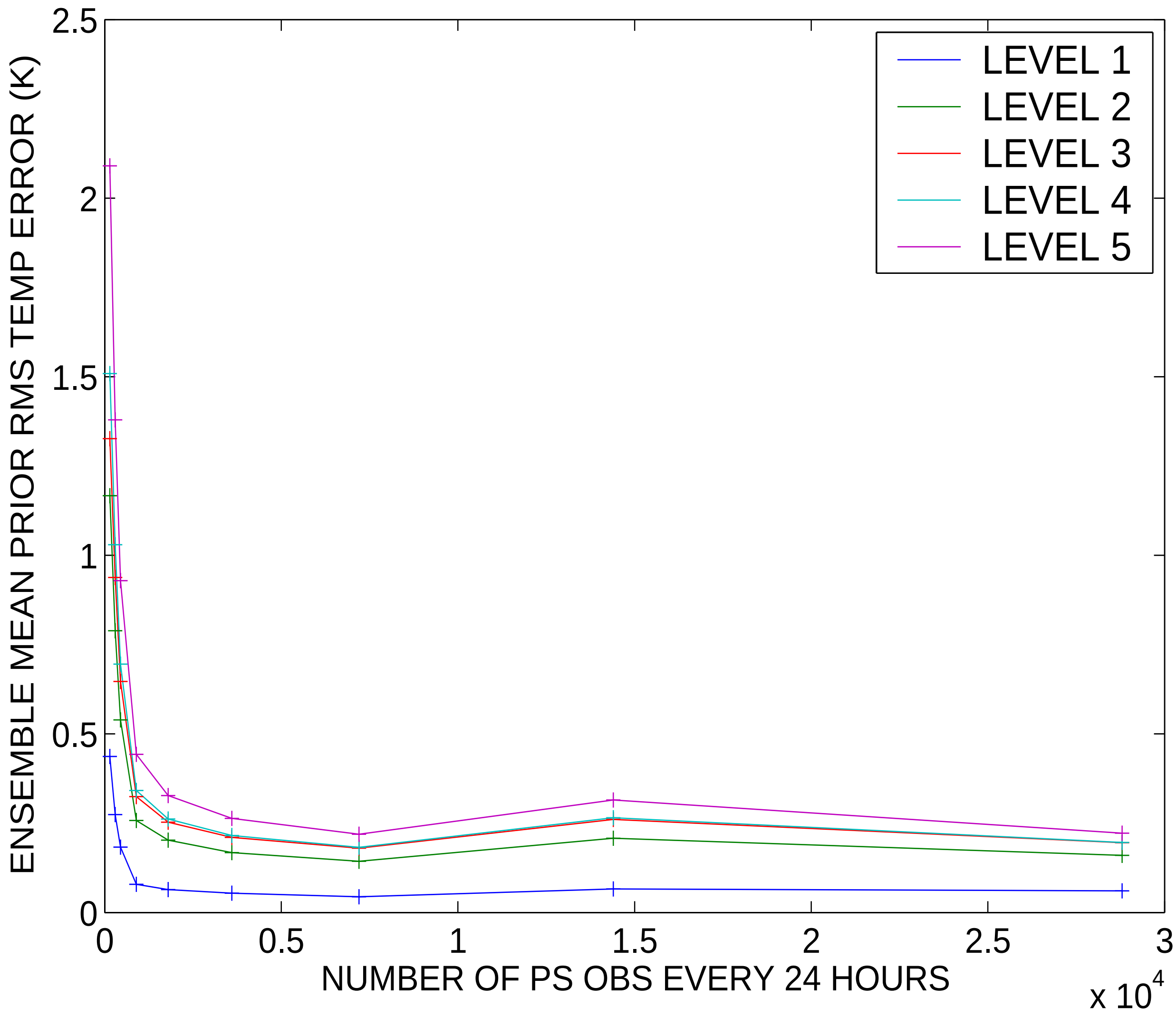
Plotting log /log of RMS shows approx. linear decrease from 150 to 7200 obs

Behavior for very large numbers of obs clearly different



Impacts of spatial density of PS obs on Temperature RMS

150, 300, 450, 900, 1800, 3600, 7200, 14400 and 28,800 every 24 hours



Behavior for Temperature (and U, V not shown) similar to that for PS
Best results for 7200 PS observations
Interior level mean T RMS of about 0.25 K for best case

Impacts of frequency of PS obs

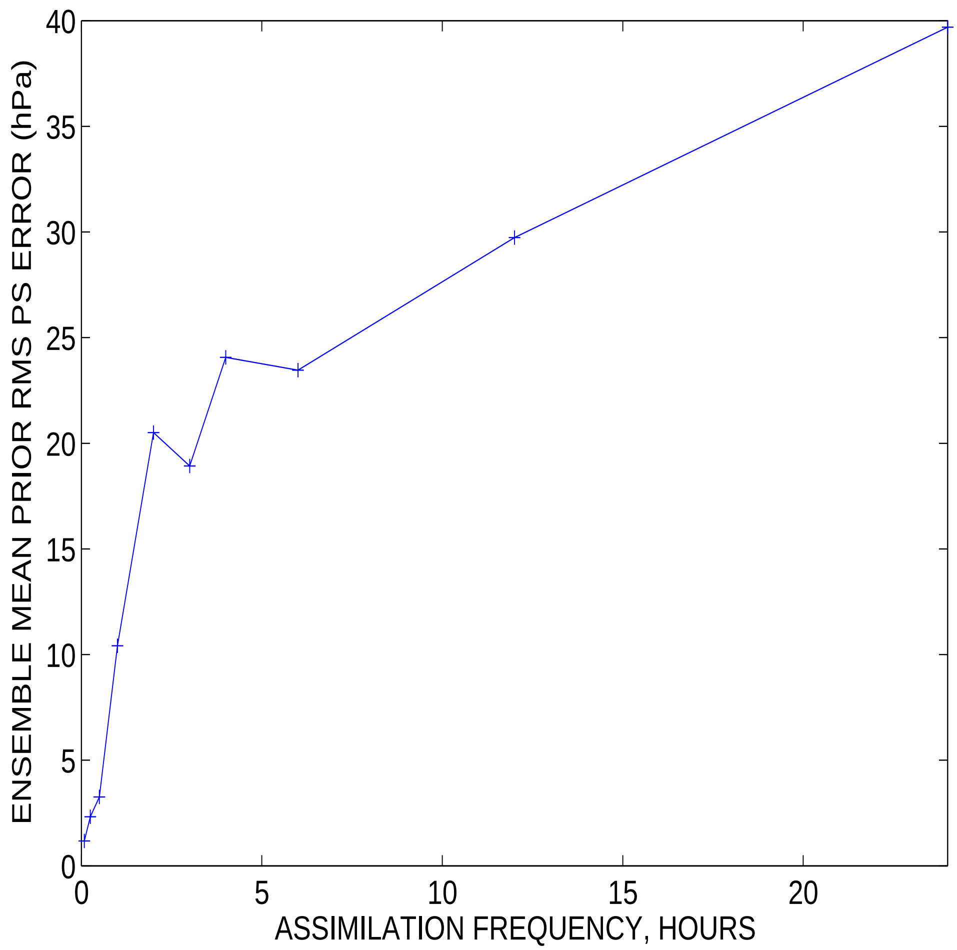
24, 12, 6, 4, 3, 2, 1 hours; 30, 15, 5 minutes; 1800 obs.

Steady RMS decrease as frequency increases

Much smaller RMS than for high density low frequency obs

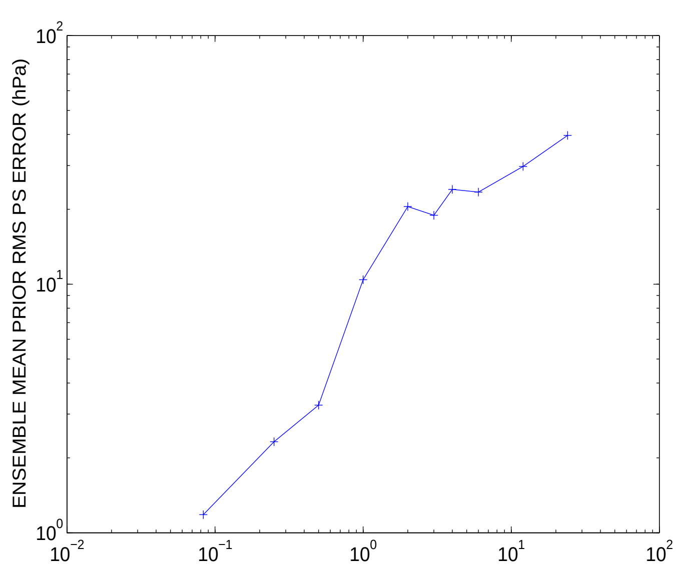
RMS < 0.02mb for 5 minutes

Strange behavior between 1 and 6 hour frequency



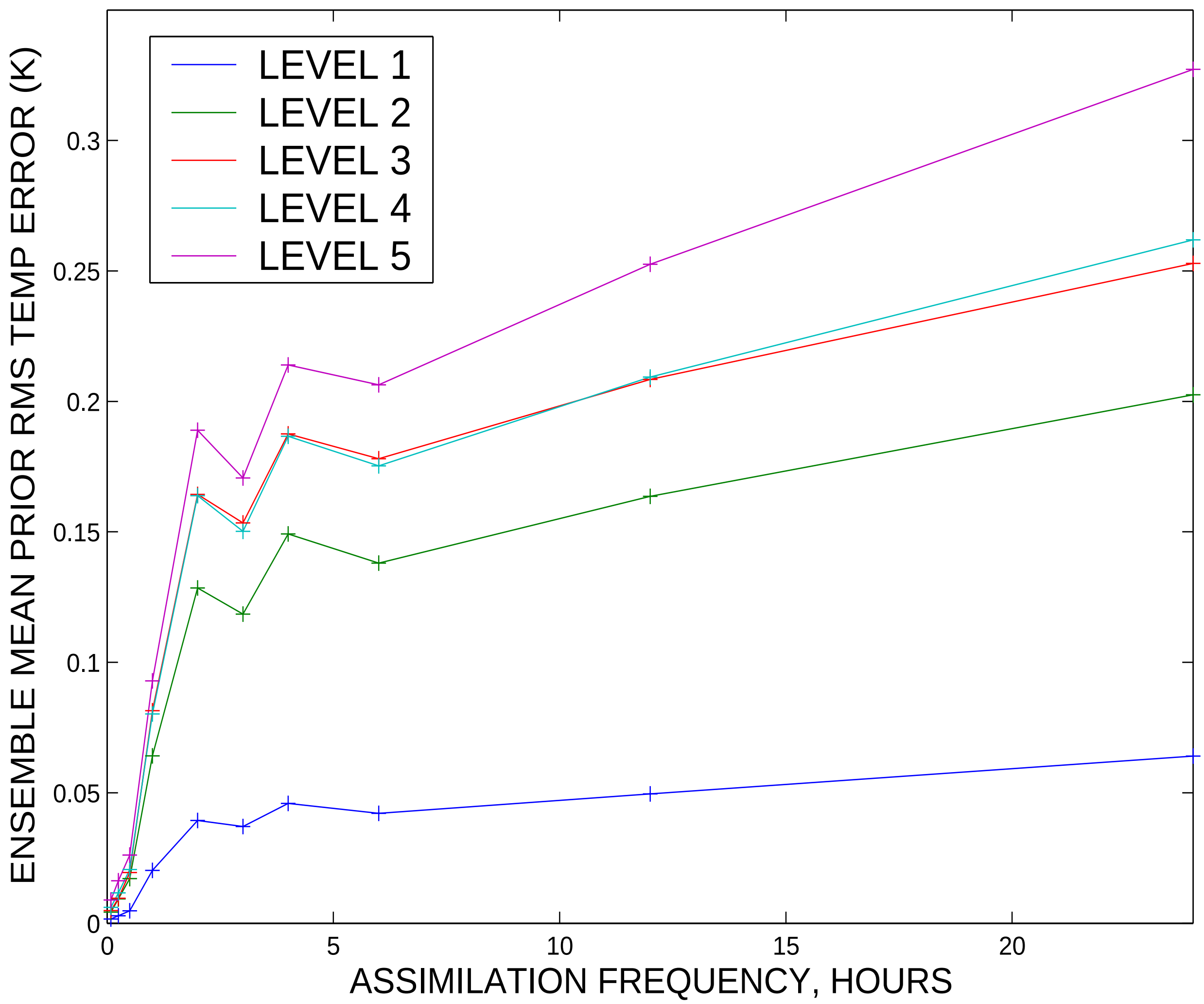
Plotting log /log of RMS shows approx. linear increase with a bump

What’s going on in the middle?



Impacts of frequency of PS obs

24, 12, 6, 4, 3, 2, 1 hours; 30, 15, 5 minutes; 1800 obs.



Temperature (and U and V, not shown) similar to PS
Consistent decrease in RMS with increased obs frequency
Errors at 5 minute frequency less than 0.01 K !!!
How low can you go?

What's going on at moderate obs frequencies?

Equilibrated model has very low gravity wave amplitude
When perturbed, 'off-attractor' gravity waves can result
Noise in observations can project off attractor

Ensemble members pulled in same direction; get phased gravity waves

Gravity wave period varies: approximately 4 hours
Gravity waves heavily damped; quickly reduced in amplitude

Low frequency (> 12 hours): gravity waves damped before next obs time

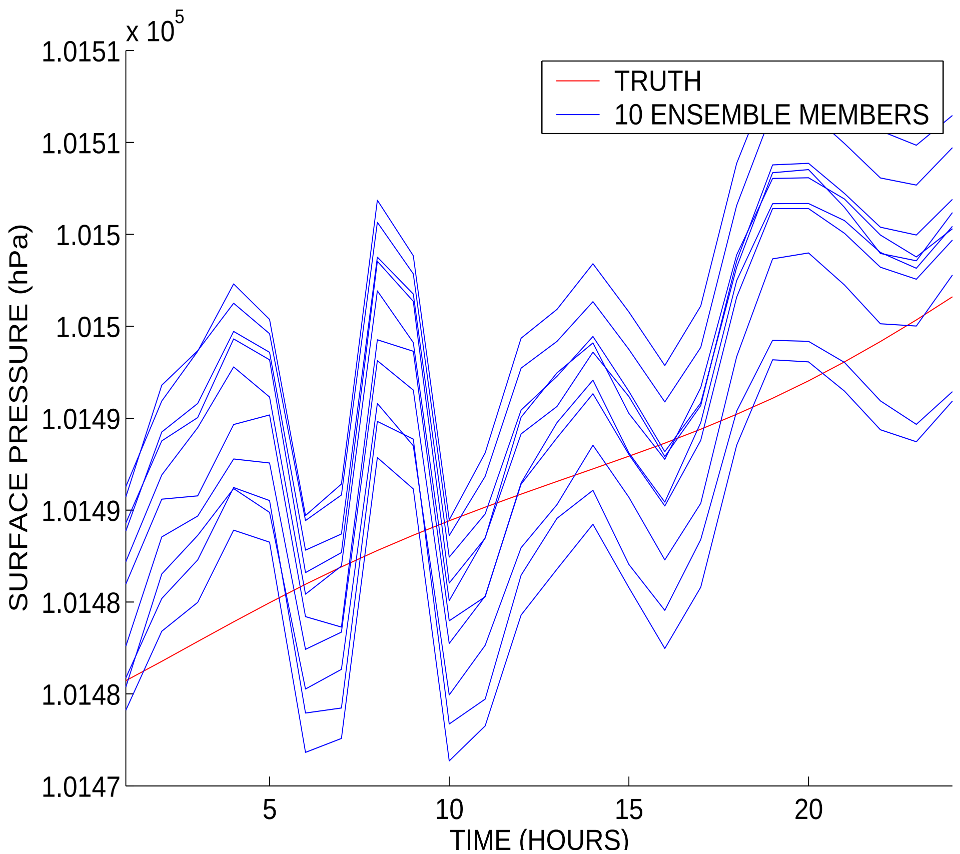
High frequency (< 1 hour): enough obs per period to control amplitude

Moderate frequency (~ 4 hours): get phased gravity waves in ensemble;
large bias; increased assimilation error

Time series of 10 out of 20 ensemble members at mid-latitude PS point.

Forecast initiated from end of 1800 PS obs every 4 hours

(Extreme example)



Why does increasing frequency do more than increasing density?

>>1. Temporal has more ‘independent’ correlation estimates

Can better eliminate sampling noise

>>2. Temporal sees observations at more ‘phases’ of wavelike structures

>>3. Large ensemble size could help to distinguish this by reducing sampling noise

These are yet to be done

>>4. Historically, high frequency obs were hard to acquire

Modern technology changes this

Exploring use of high frequency obs is planned

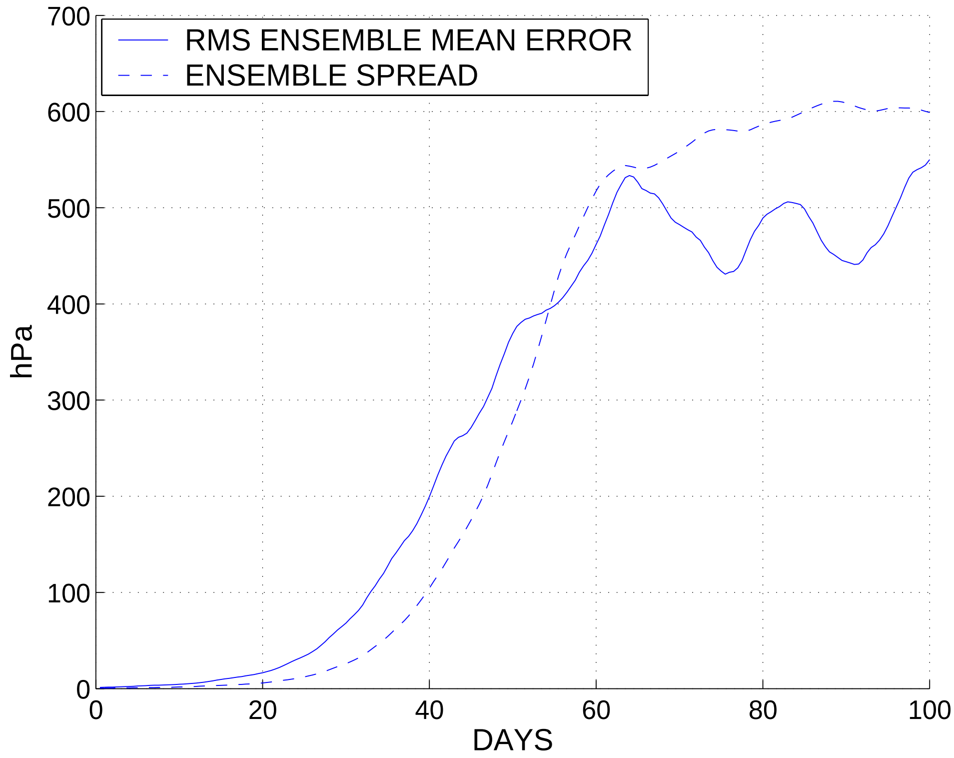
Need to demonstrate model has error growth

Free integration (forecast) at end of 1800 PS obs every 5 minutes

Error saturates at climatological values after about 60 days

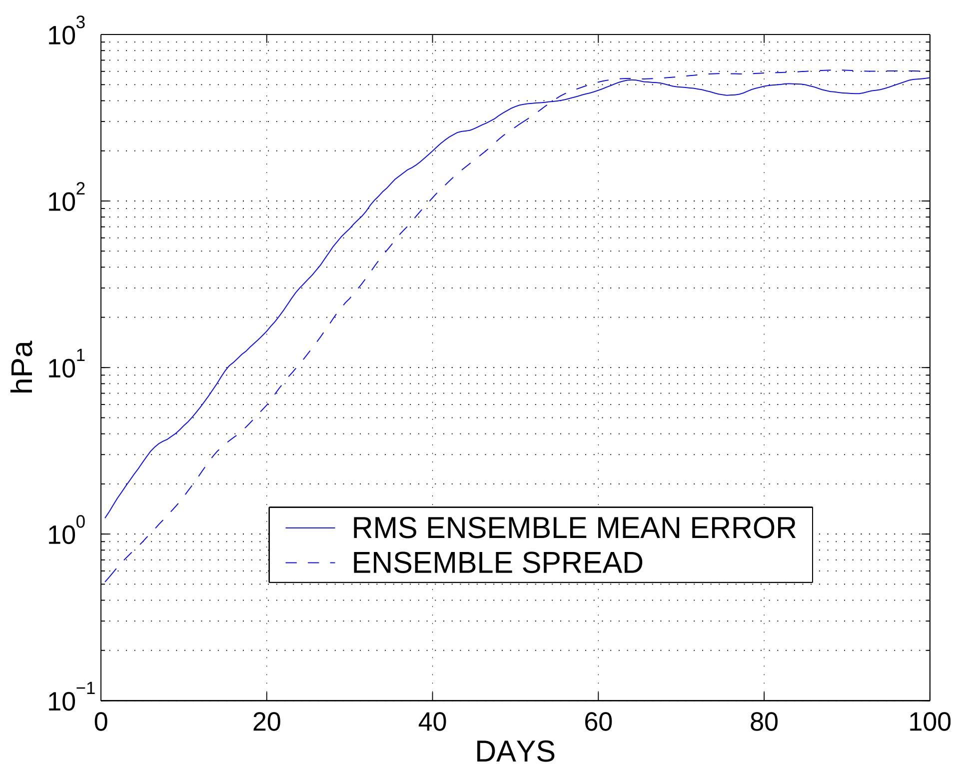
Error doubling time about 7 days

Considerably slower than real atmosphere



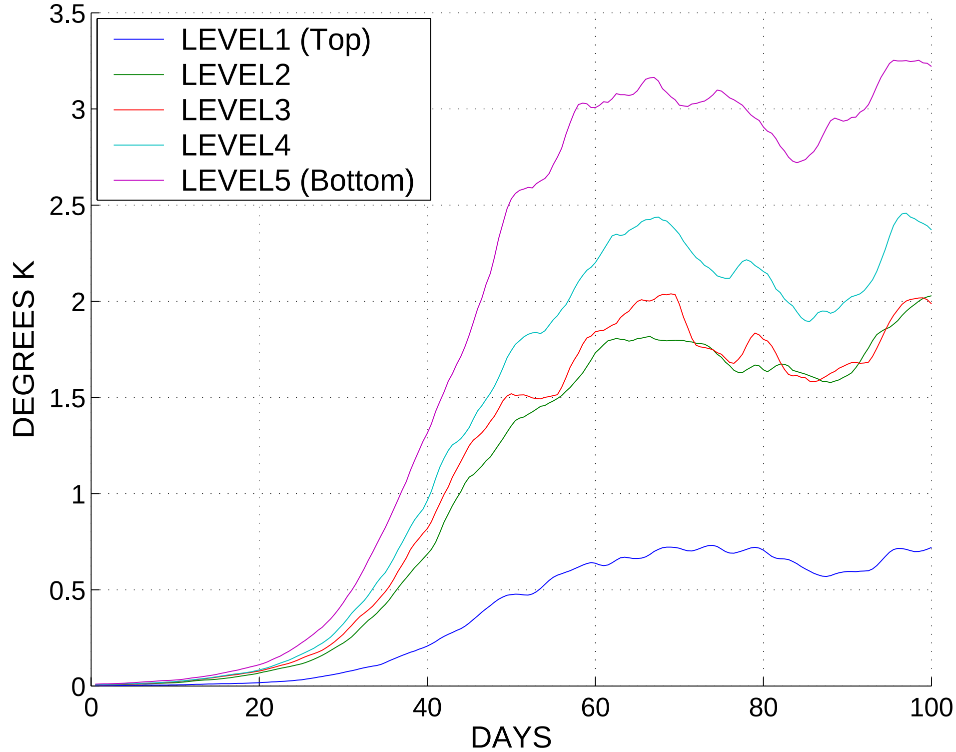
Error growth very nearly linear in log plot

Growth is almost purely exponential to saturation

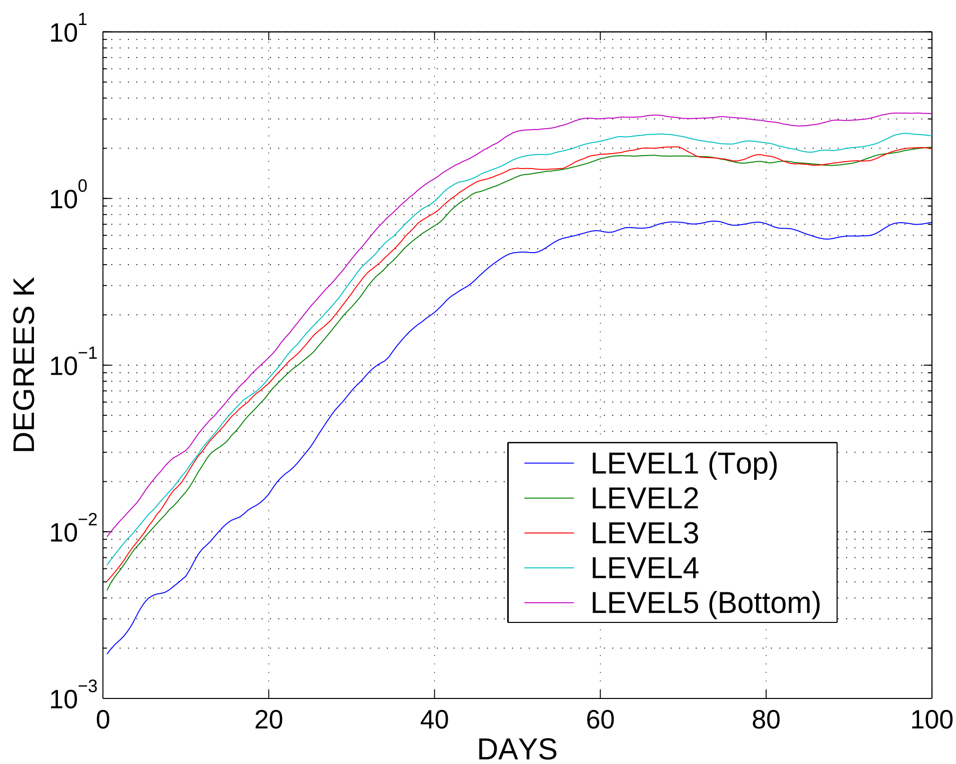


Error growth of other fields similar to PS

T (and U and V not shown) also saturate at about 60 days



Growth is very nearly exponential throughout

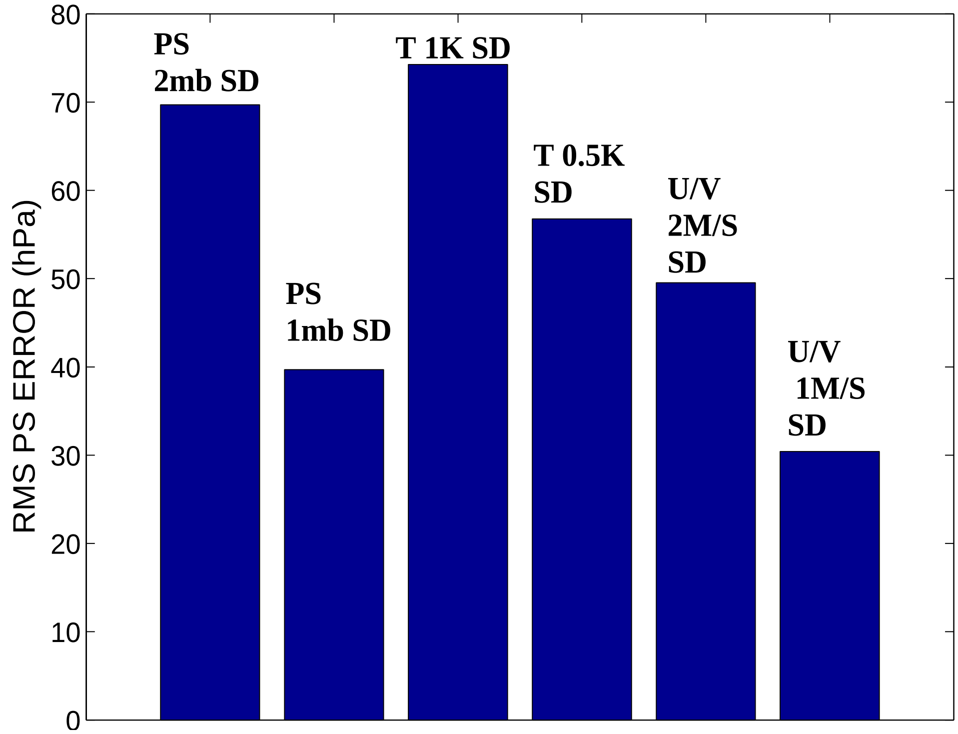


Relative Information Content of Various Surface Obs

Compare PS with T and U/V obs from lowest level

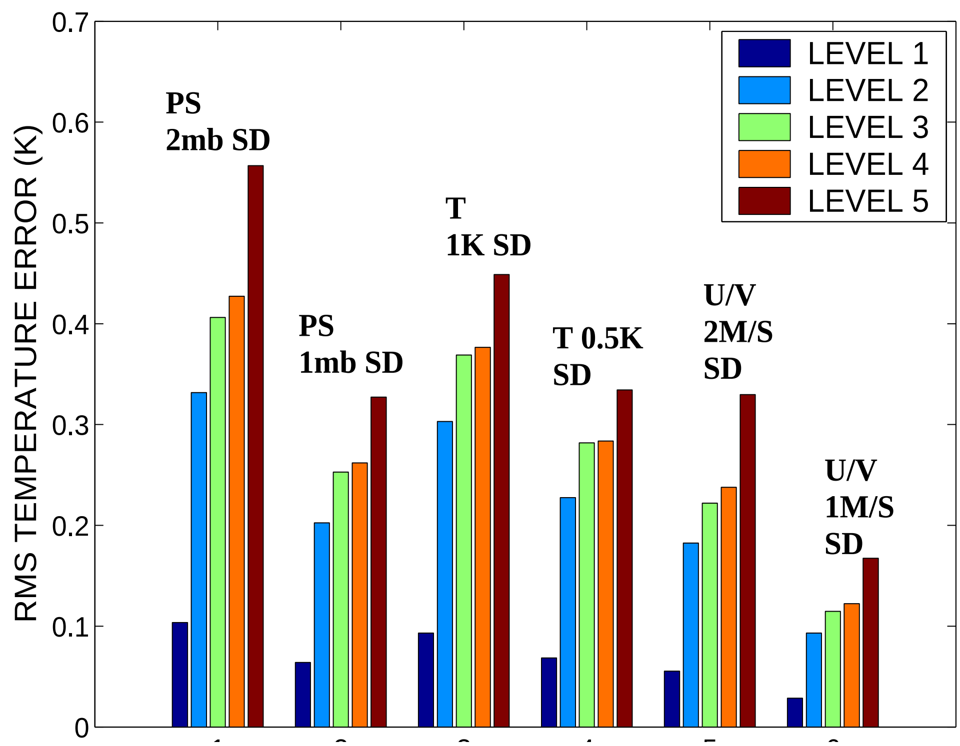
RMS error of PS prior assimilation when assimilating 1800 PS, T, or U and V wind components every 24 hours.

Two specified error SDs are checked for each.



RMS error of T for same cases.

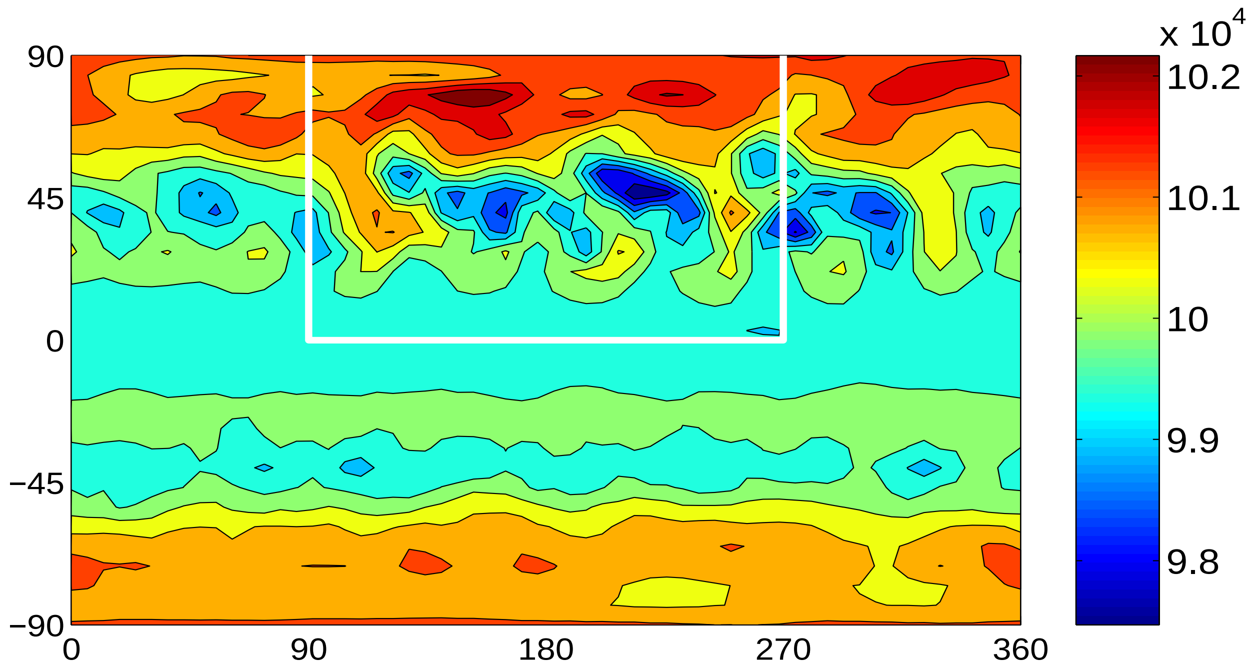
Very roughly, U/V obs with 2M/S SD have same information content as PS with 1mb SD or T with 0.5K SD.



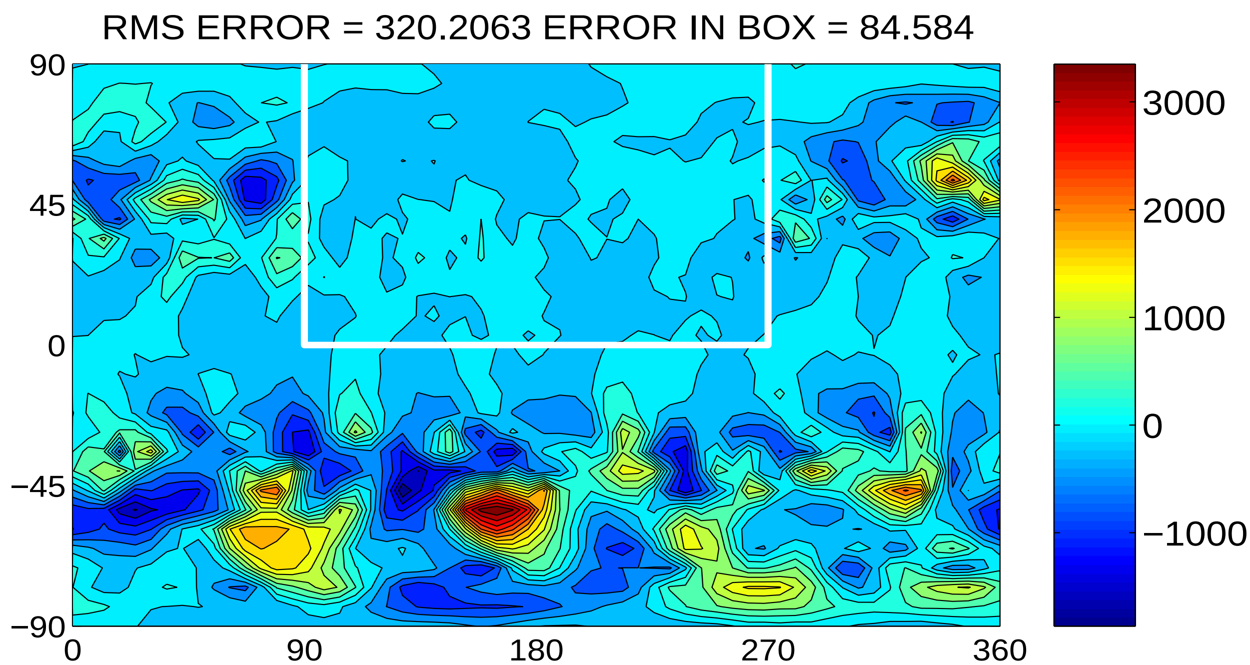
Assimilating PS over limited domains

450 PS obs every 24 hours over 1/4 of surface

Ensemble mean prior assimilation for PS at 400 days
Approaches zonal climatology with no obs information

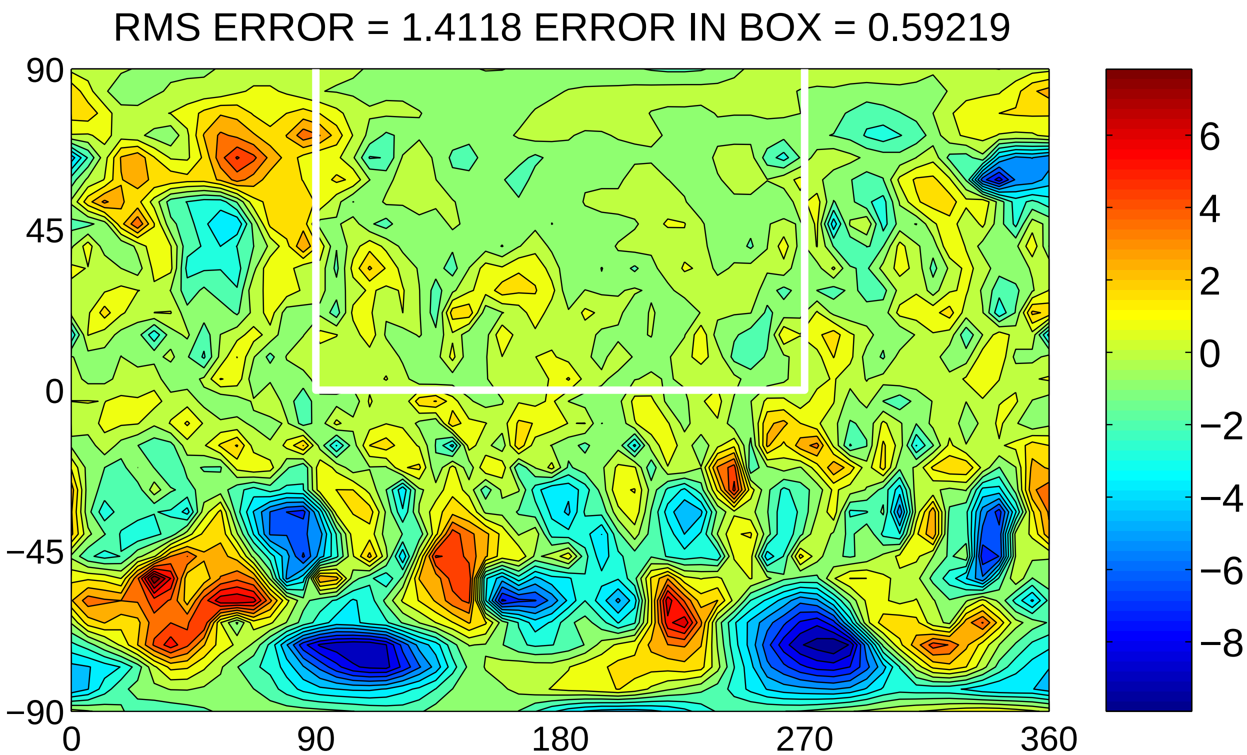


RMS Error for PS at 400 days
Error in box about twice the value for 1800 global obs



Assimilating PS over limited domain

RMS error for T at day 400;



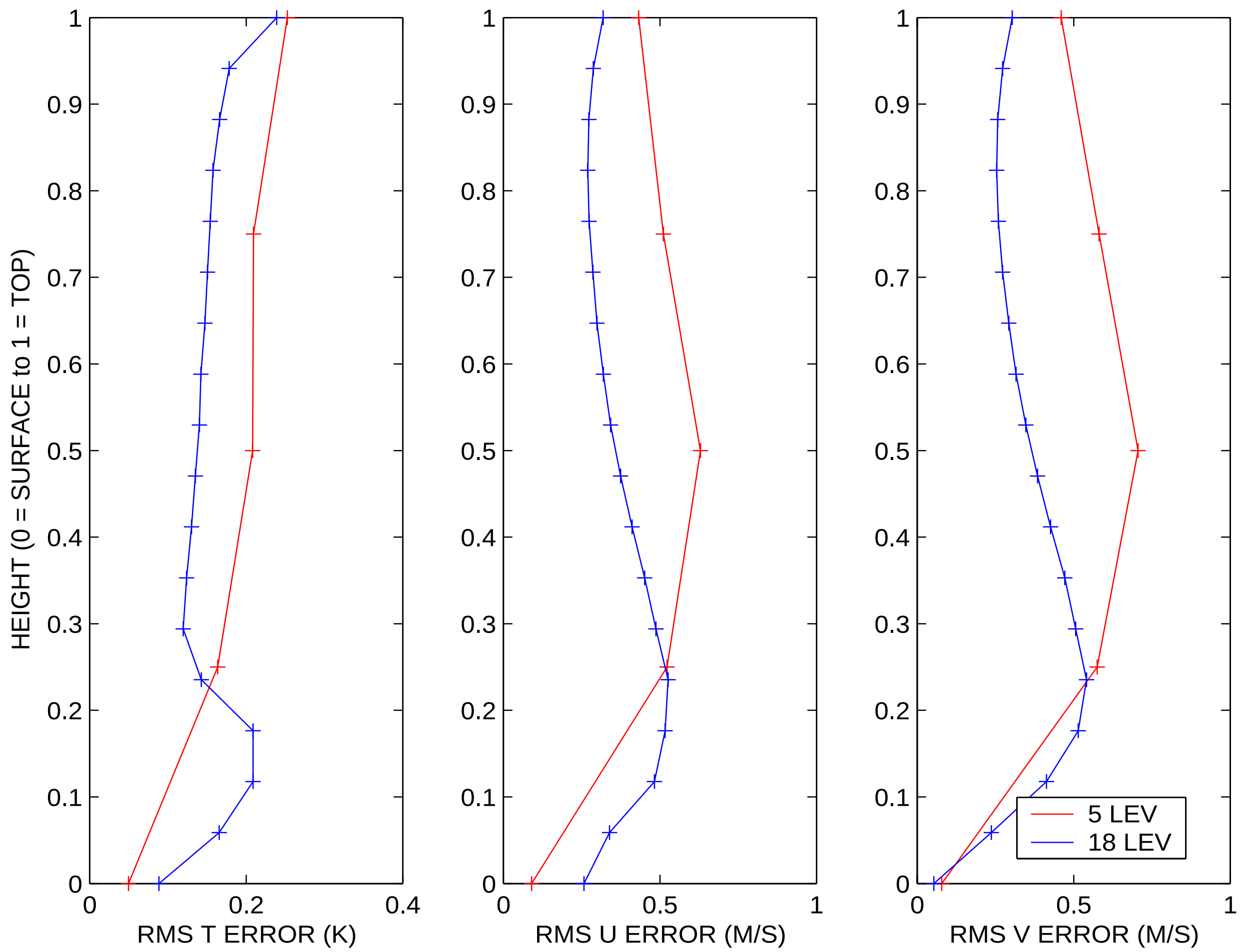
Error in box about twice that for 1800 global obs

Information is advected out of the box (to the east in mid-latitudes)

Method handles low information propagating in from upstream

Implications for regional and nested model filter data assimilation

What happens with increased resolution?



Comparison of 1800 PS obs for 5 and 18 level model
Tricky comparison, diffusion, etc. are identical
Error in upper levels of 18-level actually less

Horizontal resolution, water vapor, and more comprehensive physics:
First results in NCAR CAM at 2 degree resolution appear consistent
Results by Whitaker and Hamill with PS obs in NCEP model are good

Predictability and stochastic sub-grid scale parameterizations

Models don’t resolve all spatial scales and processes

Normally parameterized (usually by column physics)

In prediction models, physics is usually deterministic

In reality, best we can hope for is to know probability distribution for impact of unresolved processes

Can simulate this in perfect model by adding random noise to model

Here, add noise factor to temperature tendency computation

At each gridpoint, let dT/dt = MODEL * (1 + N(0, R))
N(0, R) is random number with mean 0 and standard deviation R

Independent noise at each point in current implementation

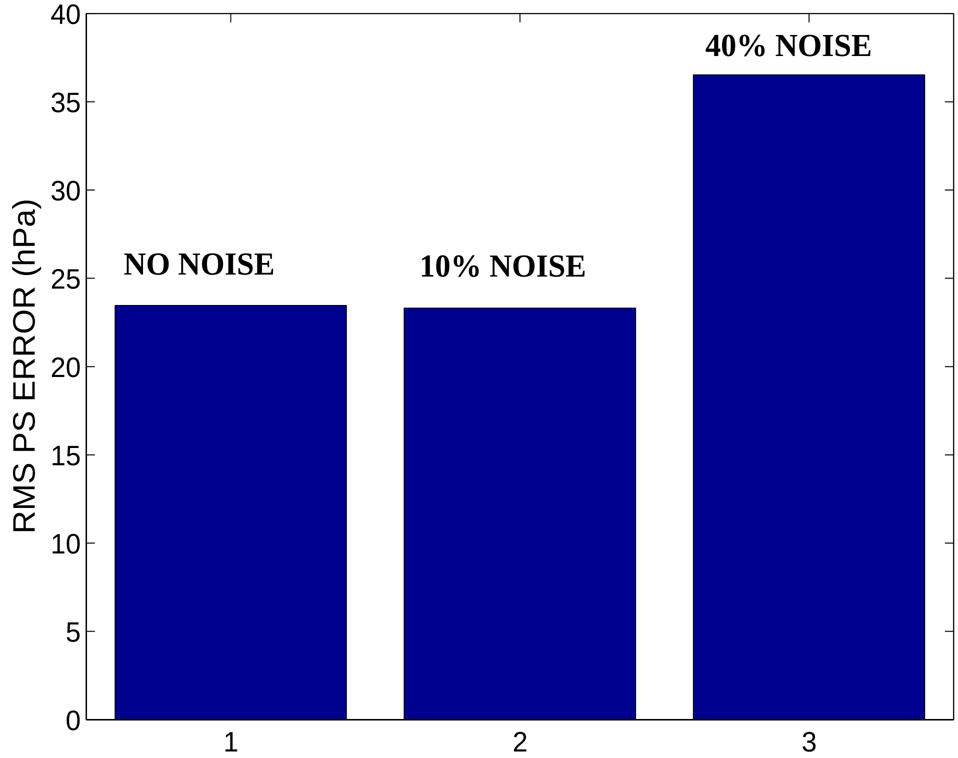
Ran cases with R = 0.1, 0.4

1800 PS obs every 6 hours (moderate gravity wave amplitude)

Impacts of sub-grid noise on Assimilation Error

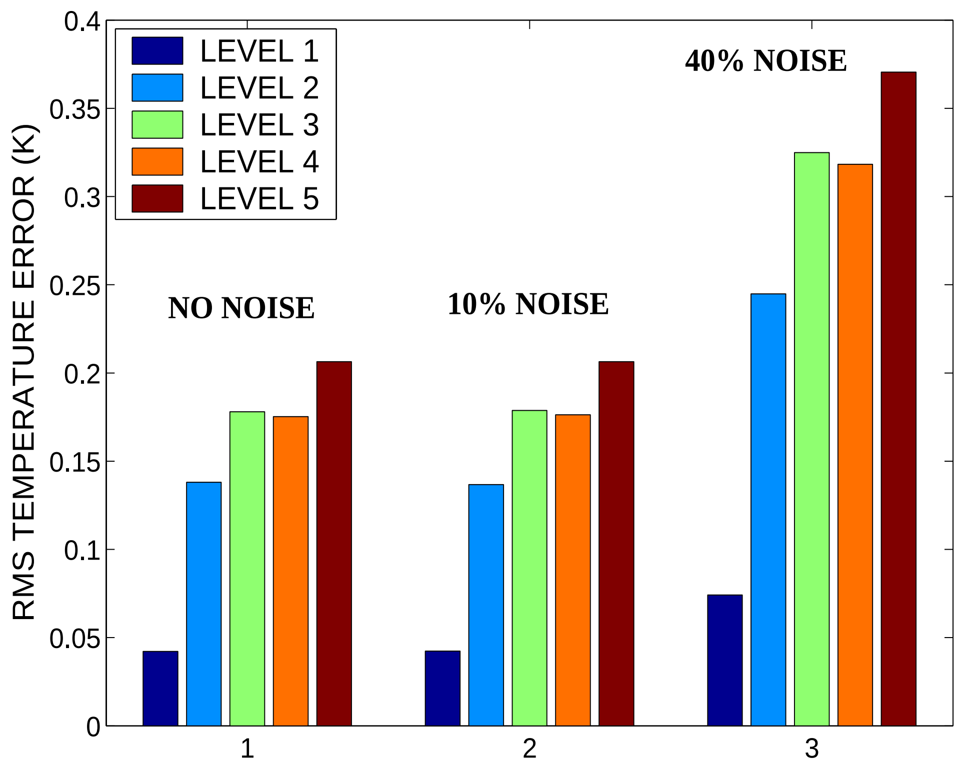
For 10%, error is slightly reduced

Adding noise in right proportions can solve gravity wave problem?



T results similar.

40% case has much increased error.

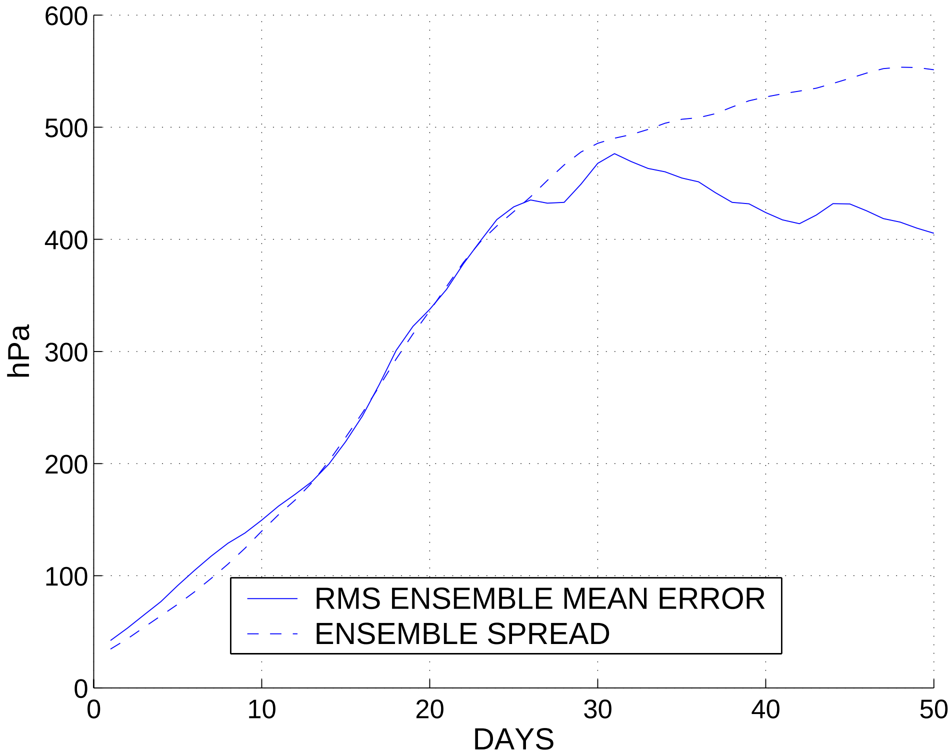


Error growth and predictability with sub-grid scale noise

PS error growth is mostly linear at first.

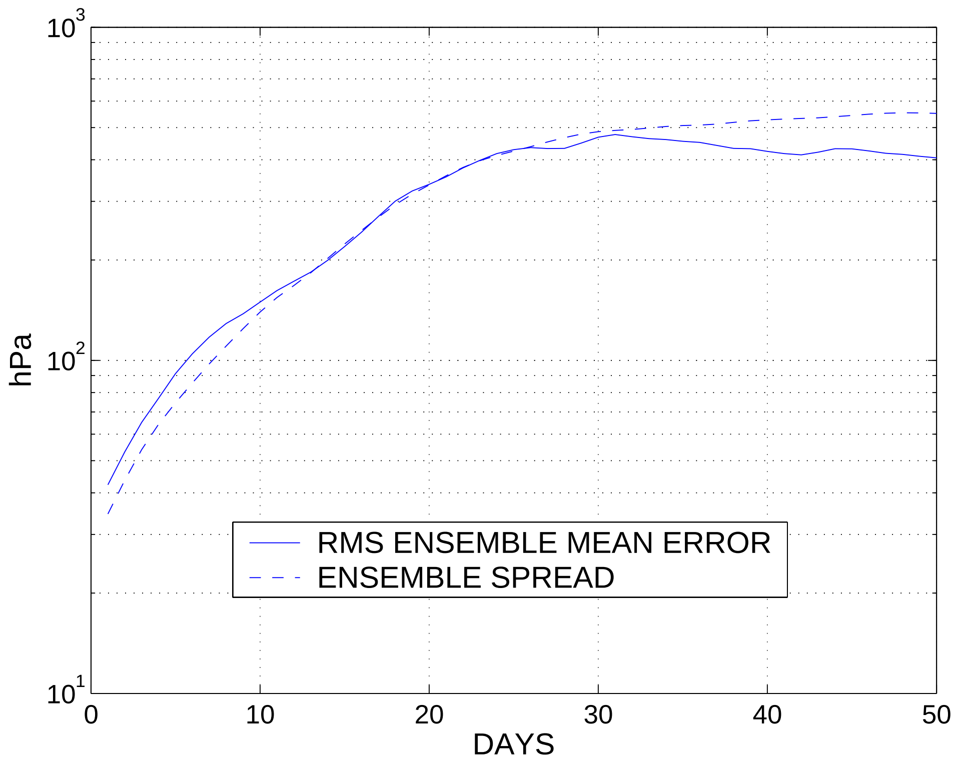
Then hint of exponential after day 15?

Saturates by day 25.



Should expect error growth in real systems to look like this.

Operational systems do not at this point?



Conclusions and future work

1. (Small) ensemble filter can extract lots of information

2. Increasing temporal density of obs may be very effective

3. In perfect world, surface obs deliver accurate assimilations

4. What can high frequency surface obs do in real assimilation / prediction problems?

5. Bias, bias, and bias are key remaining problems

6. Predictability studies must be done in assimilation / prediction context with stochastic sub-grid scale parameterizations

Next step: Moderate resolution GCM with physics:
(B-Grid I release?, CAM?, NCEP?)

Dealing with bias in ensembles is remaining problem

Bayesian Theory supporting filters excludes bias

But, we know there are many violations of the Gaussian assumptions we make for implementation

Need to build an additional a priori model of bias

Covariance inflation and related tricks are one simple model
- Have some advantage by retaining correlation structure
- Simply States that there is an additional Gaussian component of error that is not accounted for by the model

Can we do more sophisticated, adaptive models?

With ensemble and known observation error distribution, can determine expected value of sum of model and observation bias for any observation

In other words, is the distance between the prior obs estimates and the obs inconsistent?

Can aggregate these statistics in time, or space or both

Need to partition unaccounted error into one of three bins:
1. Model first moment bias (error)
2. Model second moment bias (error)
3. Observation bias (error)

Dealing with bias in ensemble filters (cont.)

May be easy to partition between 3 and combined (1, 2)
- Similar to buddy checks
- Are observations in same ‘area’ not consistently inconsistent
- If so, much more inconsistent obs should have large bias

Tricky problem, how to partition bias between first and second moment in model

If it’s first moment, just let observation be more compelling

If it’s second, need to reduce decrement in spread

Initial results playing with this have been very successful in very large bias systems

Need to try out in a real setting

Note: this should eventually replace a part of quality control