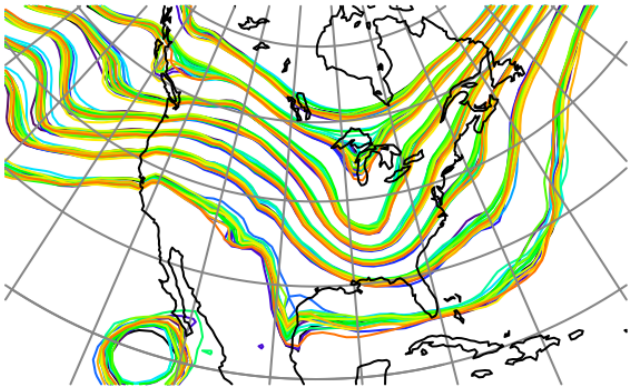
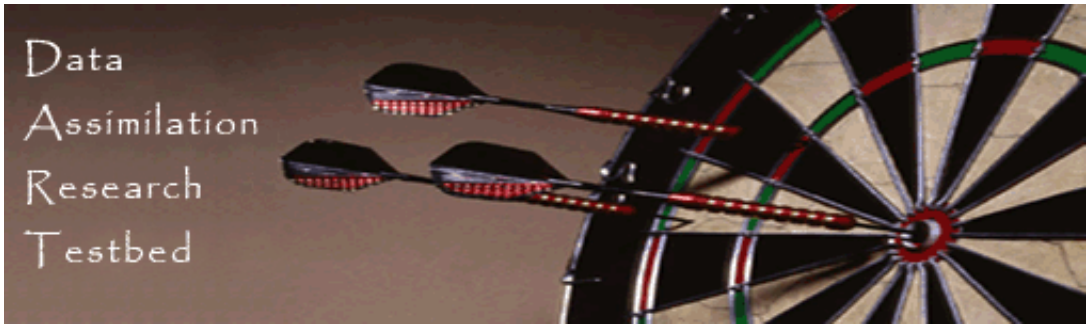


DA and Short-term Prediction Motivate the Need for Careful Software Development In Earth System Models: A Case Study

Jeff Anderson, NCAR Data Assimilation Research Section



A Story About a Finite Volume Atmospheric Dynamical Core

Core originally developed at NASA.
Adopted by GFDL and NCAR for climate investigations.
Simulated at least $O(10^5)$ years, maybe more.

My team added an ensemble DA capability in 2006.

That's when we found the following:

Diagnosis of Noise in the CAM Finite Volume core using DART

Kevin Raeder*

Jeff Anderson*

Peter Lauritzen+

Tim Hoar*

*NCAR/CISL/IMAGe/DAReS
+NCAR/ESSL/CGD/AMPS

The National Center for Atmospheric Research is sponsored by the National Science Foundation.

CAM & DART

CAM = 3.5.xx, Finite Volume core, 1.9x2.5, 30 min Δt.

DART = Data Assimilation Research Testbed, an ensemble Kalman filter data assimilation system.

Assimilate observations used in operational forecasting:

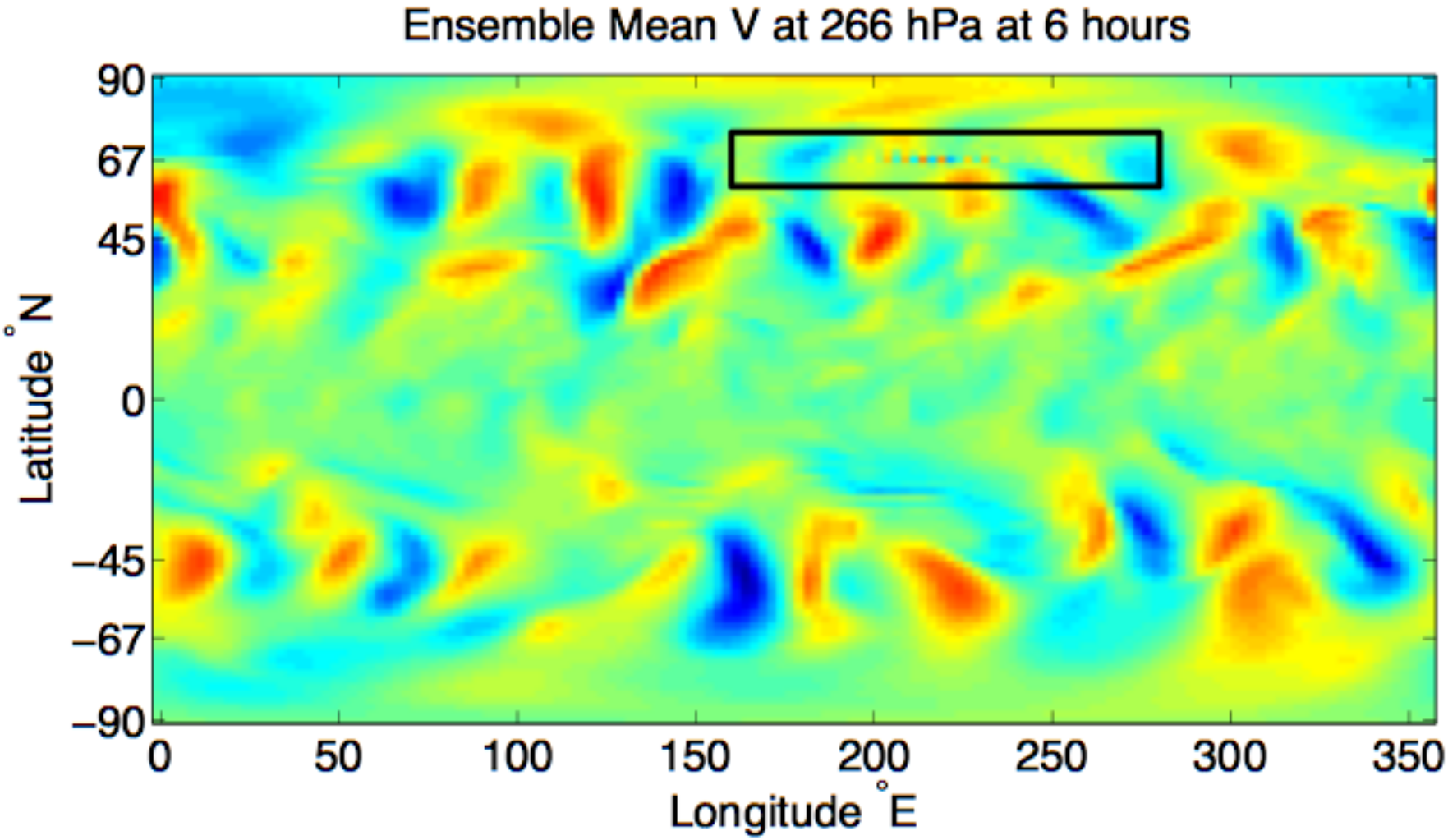
→U, V, and T from radiosondes, ACARS, and aircraft,

→U and V from satellite cloud drift winds,

→every 6 hours to bring CAM as close to the atmosphere as possible, balancing the obs and model errors.

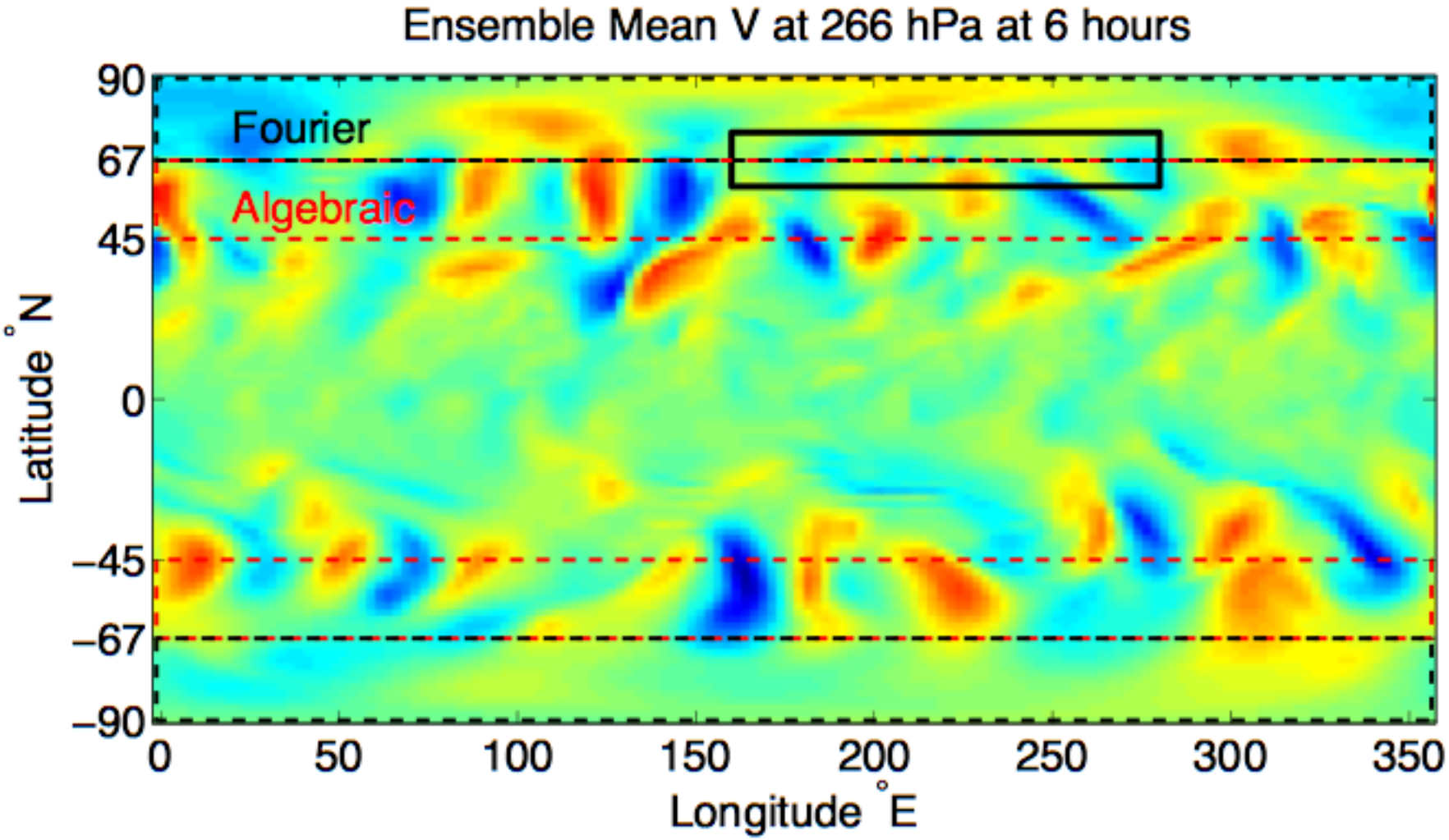
This system is competitive with operational weather centers' data assimilation systems.

“Houston, we have a Problem.”



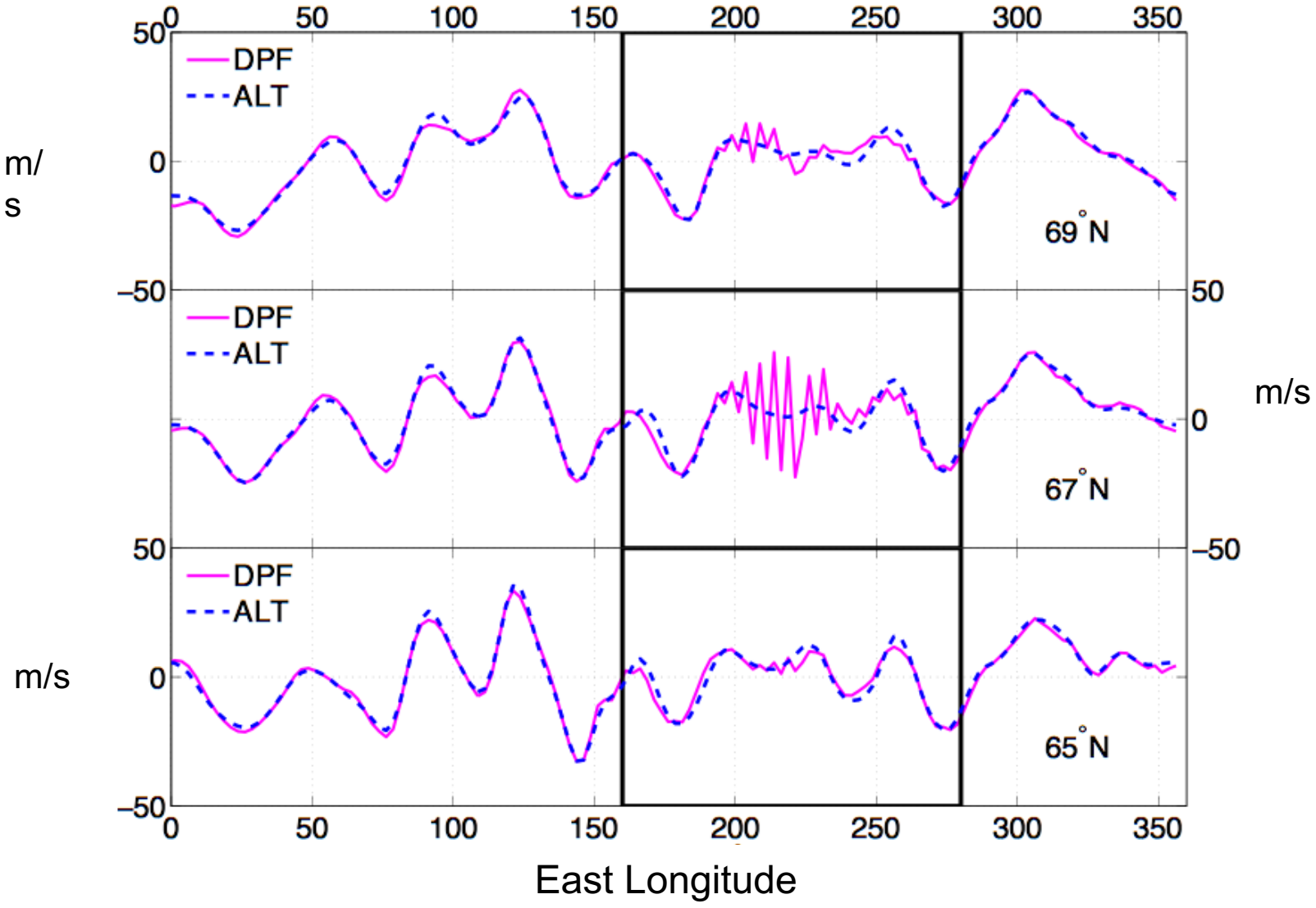
CAM FV core - 80 member mean - 00Z 25 September 2006

Suspicions turned to the polar filter (DPF)



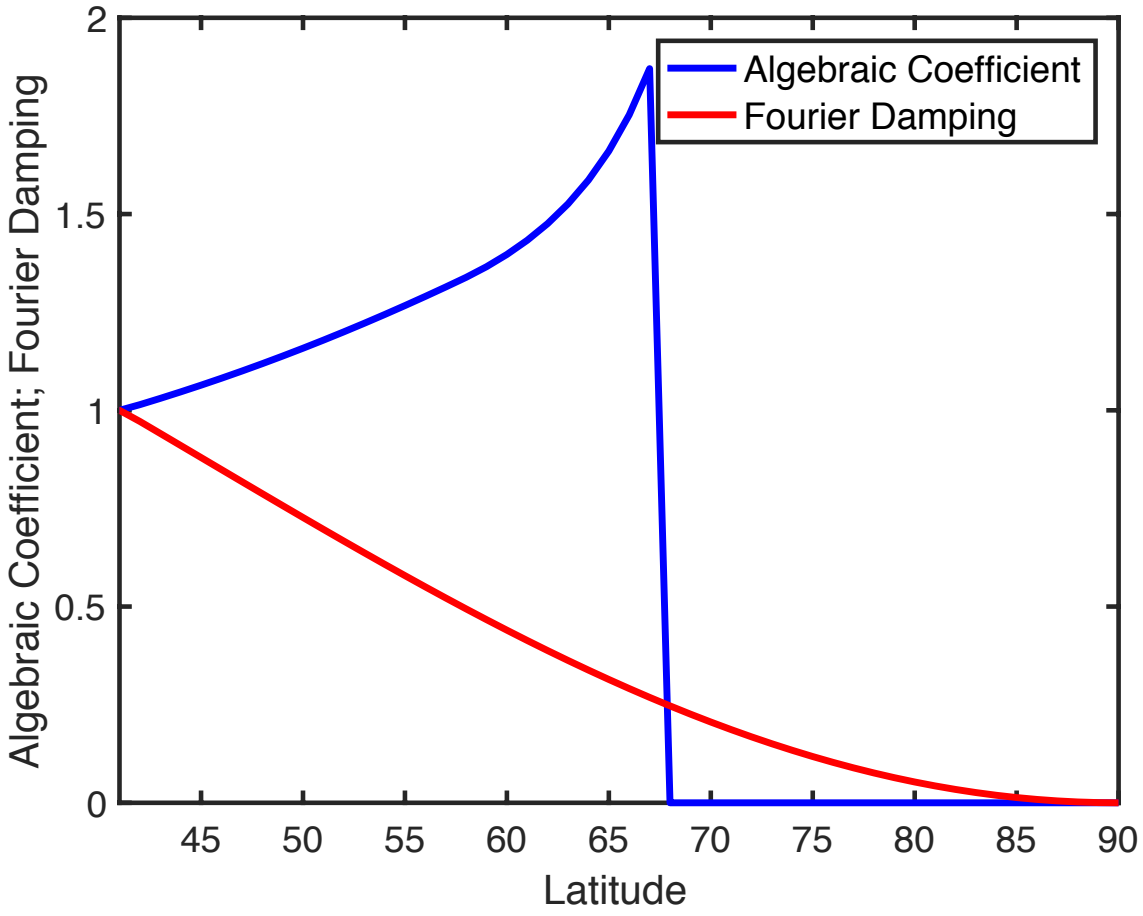
CAM FV core - 80 member mean - 00Z 25 September 2006

Three adjacent E-W cross-sections from the region of the discontinuity reveal more detail.



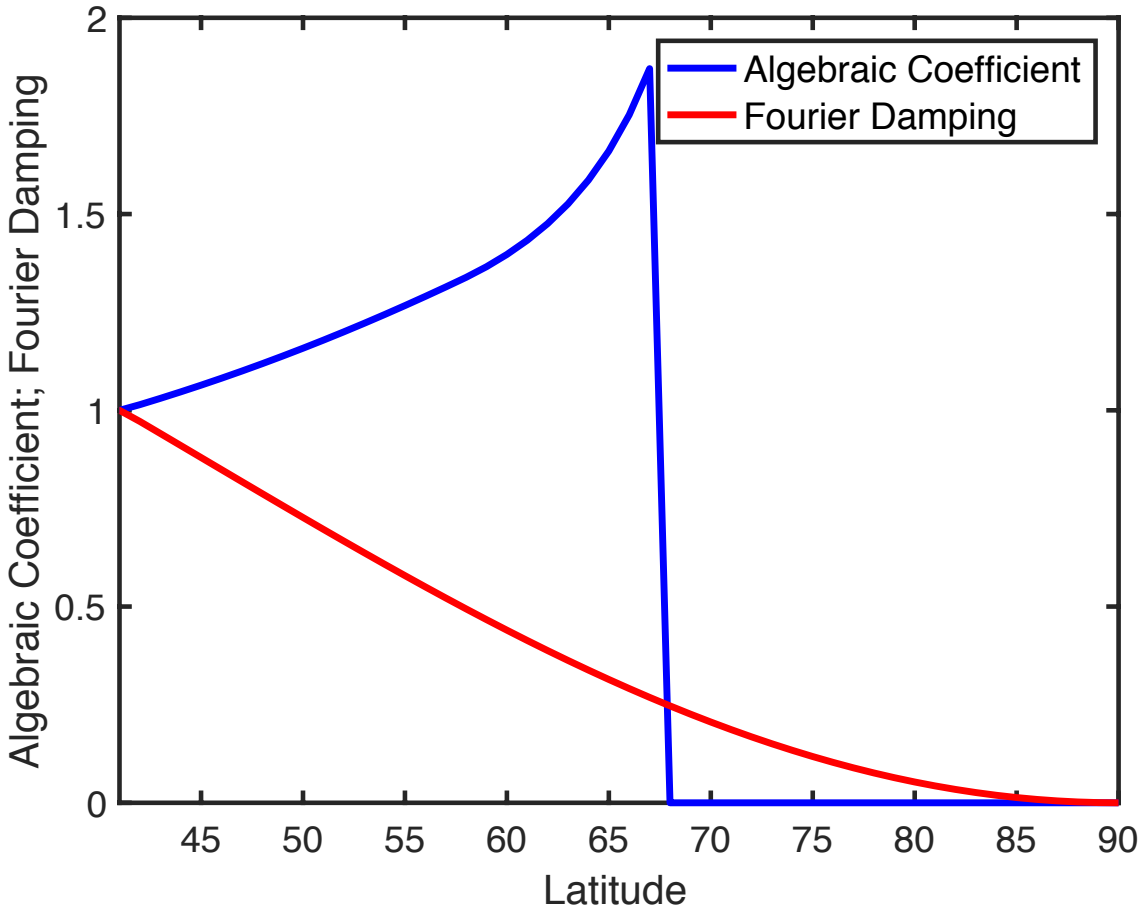
Ensemble Mean V @ 266hPa - 00Z 25 Sep 2006 - CAM FV core

Combination of algebraic filter and polar Fourier filter



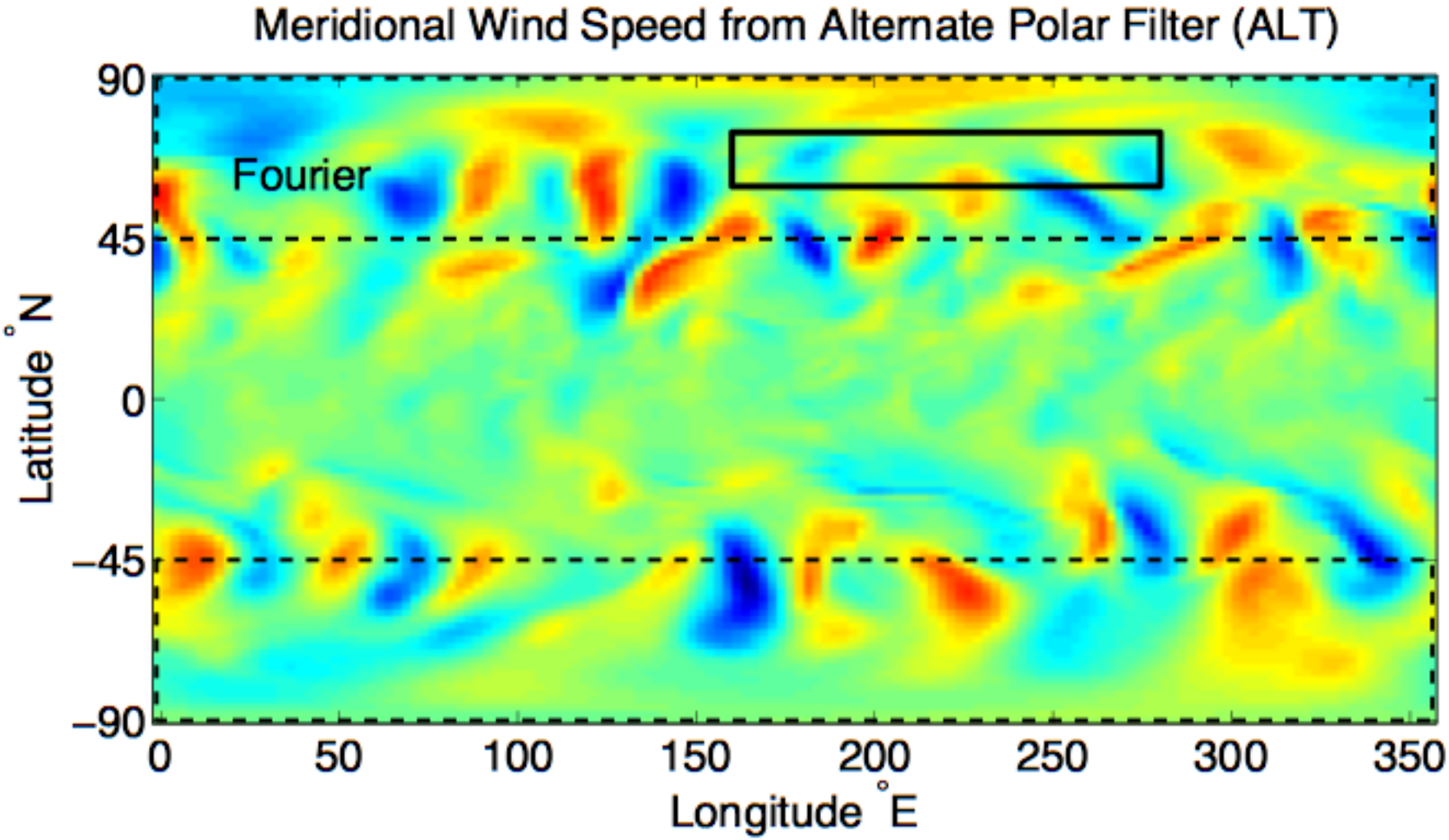
Original: Both fourier and algebraic get more dissipative between 41 and 68 degrees. Then algebraic turns off.

Combination of algebraic filter and polar Fourier filter



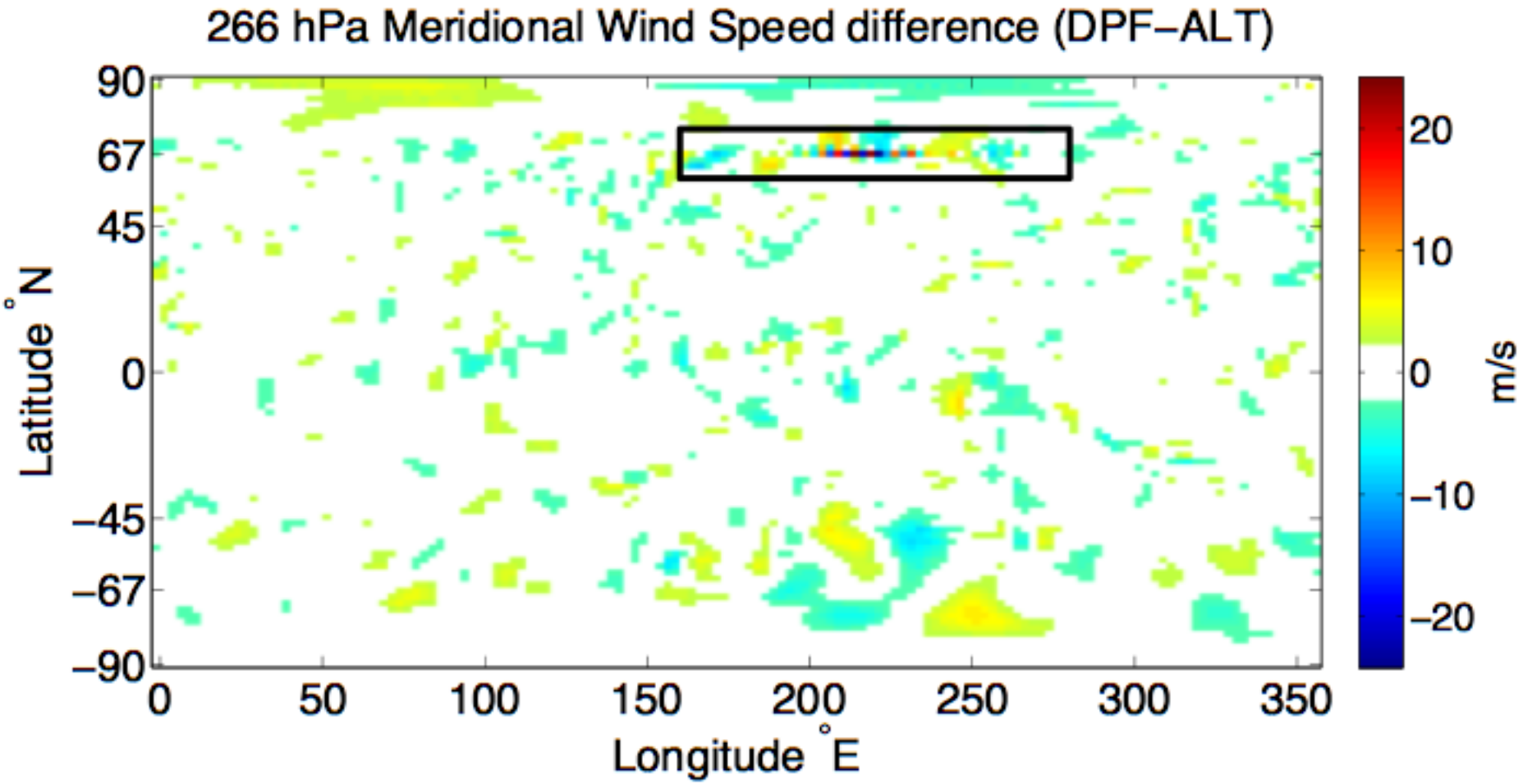
Unclear what actual intent was. Probably to turn algebraic off gradually between 41 and 68?

Using a continuous polar filter (alt-pft) does not show this effect.



Ensemble Mean V @ 266hPa - 00Z 25 Sep 2006 - CAM FV core

The differences are minimal except at the transition region of the default polar filter.



Ensemble Mean V @ 266hPa - 00Z 25 Sep 2006 - CAM FV core

That wasn't so bad!

- The use of DA diagnosed a problem that had been unrecognized (or at least undocumented).
- The problem can be seen in 'free runs' - it is not a data assimilation artifact.
- Without assimilation, can't get reproducing occurrences to diagnose.
- Could have an important effect on any physics in which meridional mixing is important.
- The alternate polar filter 'fixes' this problem.

Benefits of Ensemble DA for Model Development

- Confront model with observations, look for inconsistencies.
- Create reproducible, realistic model cases.
  - Enable study of particular events.
  - Compare to existing models for these cases.
- Sensitivity analysis, correlate any function of state with any function of state.
- Direct estimation of model parameters.

Model Design Requirements for Ensemble DA

Results from NCAR Singletrack project,
Working group on Data Assimilation.

Model Design Requirements for Ensemble DA

1. Can perform a sequence of short integrations with minimal computational overhead relative to a single long integration.
2. Can stop and restart exactly.
3. Precise, accessible definition of prognostic state.
4. Easy to invoke a range of damping and smoothing operators so that model is stable when using DA.
5. Ability to access and modify model parameters.
6. Compute forward operators efficiently at high frequency.
7. Tangent linear capability and adjoint model capability?
8. Tangent linear and adjoint for forward operators?

Software Engineering for Living Science Codes

This code was used at 3 major institutions.
At least O(10^5) years of model integration.
It is obviously incorrect.
Leads to clearly visible concerns in instantaneous fields.

Why didn't anybody notice?

Software Engineering for Living Science Codes

Why didn't anybody notice?

- Model doesn't crash.
- There is no first principles algorithm.
- There is no unit test available (or possible?).
- In long runs, no simple 'test' for bad behavior?
- This piece of code was used to 'tune' model behavior.
- It clearly evolved after original writing.
- The code contained negligible documentation.
- The code is difficult to interpret.
- Code path is controlled by a number of external parameters.

Why didn't anybody notice?

```
!************
! Cell center
!************
      do j=js2g0,jn2g0
           sc(j) = (coszc/cosp(j))**2

         if(sc(j) > D1_0 ) then
          if(fft_flt .eq. 0 .and. sc(j) <= D2_0) then
            sc(j) =  D1_0 +  (sc(j)-D1_0)/(sc(j)+D1_0)
          elseif(fft_flt .eq. 0 .and. sc(j) <= D4_0) then
            sc(j) =  D1_0 +  sc(j)/(D8_0-sc(j))
          else

! FFT filter
            do i=1,im/2
              phi = dl * i
              damp = min((cosp(j)/coszc)/sin(phi),D1_0)**2
              if(damp < cutoff) damp = D0_0
              dc(2*i-1,j) = damp
              dc(2*i  ,j) = damp
            enddo

          endif
         endif
       enddo
```

Software Engineering for Living Science Codes

Axioms:
Scientists aren't software engineers,
Software engineers aren't scientists,
Both fields require non-trivial expertise.

Scientists should do careful science.
Software engineers should do careful engineering.

Should work closely together during implementation.

Software Engineering for Living Science Codes

Scientists should do careful science:

Precise, reproducible, documented, justified.

SEs should do careful engineering:

All code should be clear, documented.

Separation of development from supported code.

Pair should confirm that code does what scientists want.

Software Engineering for Living Science Codes

Key points:

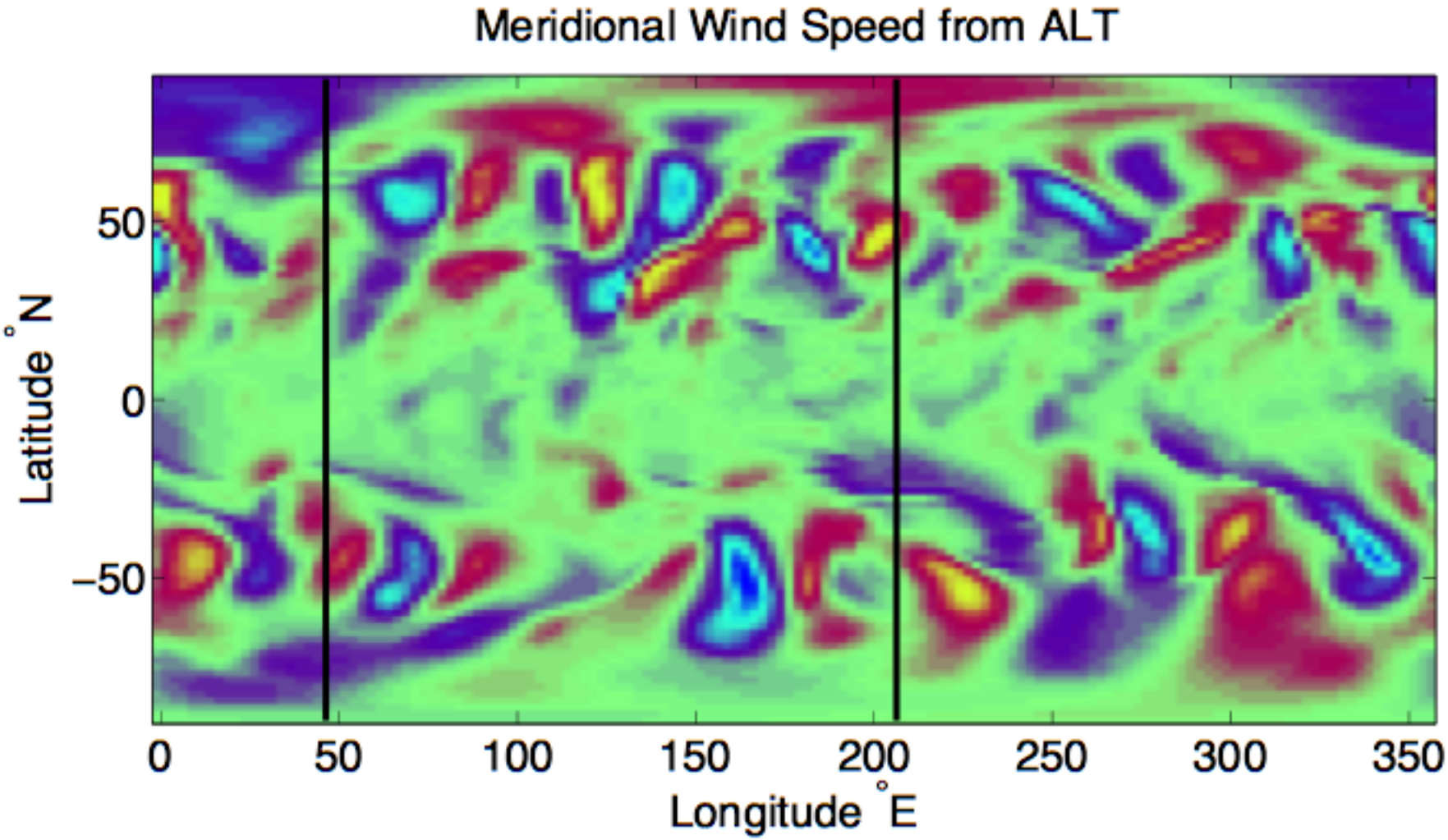
1. Highly unlikely that any existing GCM is ‘error-free’
2. How can one interpret results from something known to be erroneous.
3. Subsequent developments may be compensatory.

My opinion:

- Much more careful development is essential.
- Much more comprehensive unit testing where possible.
- Far more important than new parameterizations, etc.

More suspicious patterns, not fixed by ALT_PFT

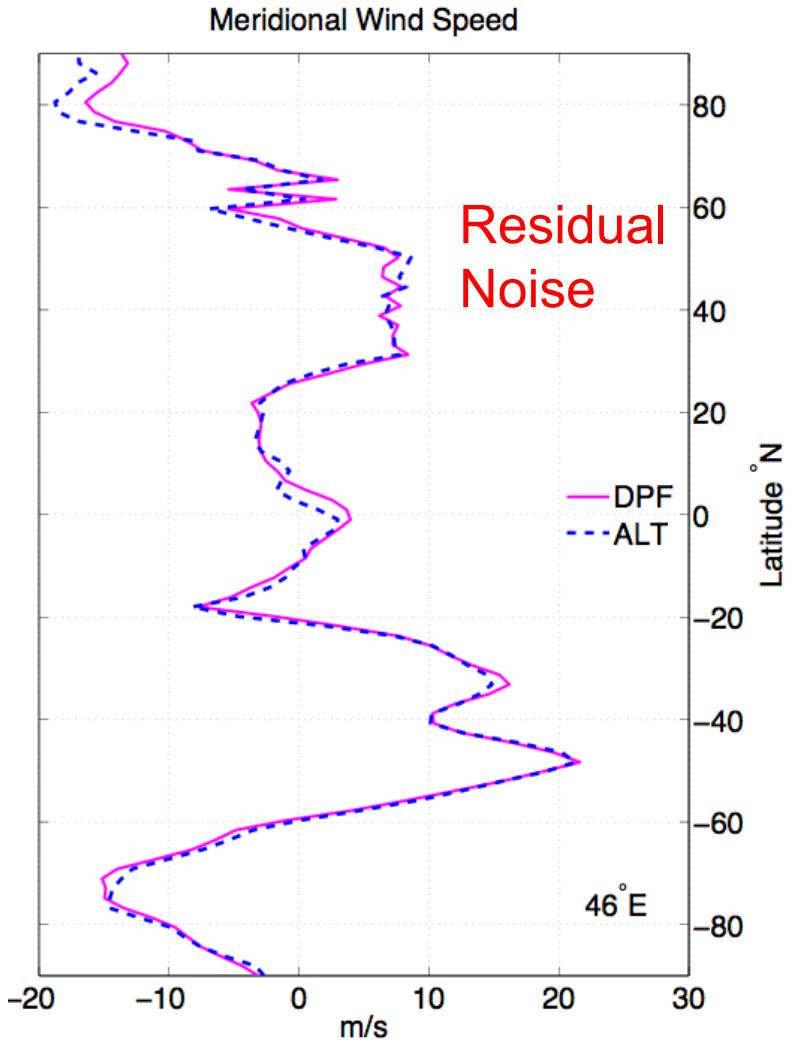
2 Δy noise in ensemble average V



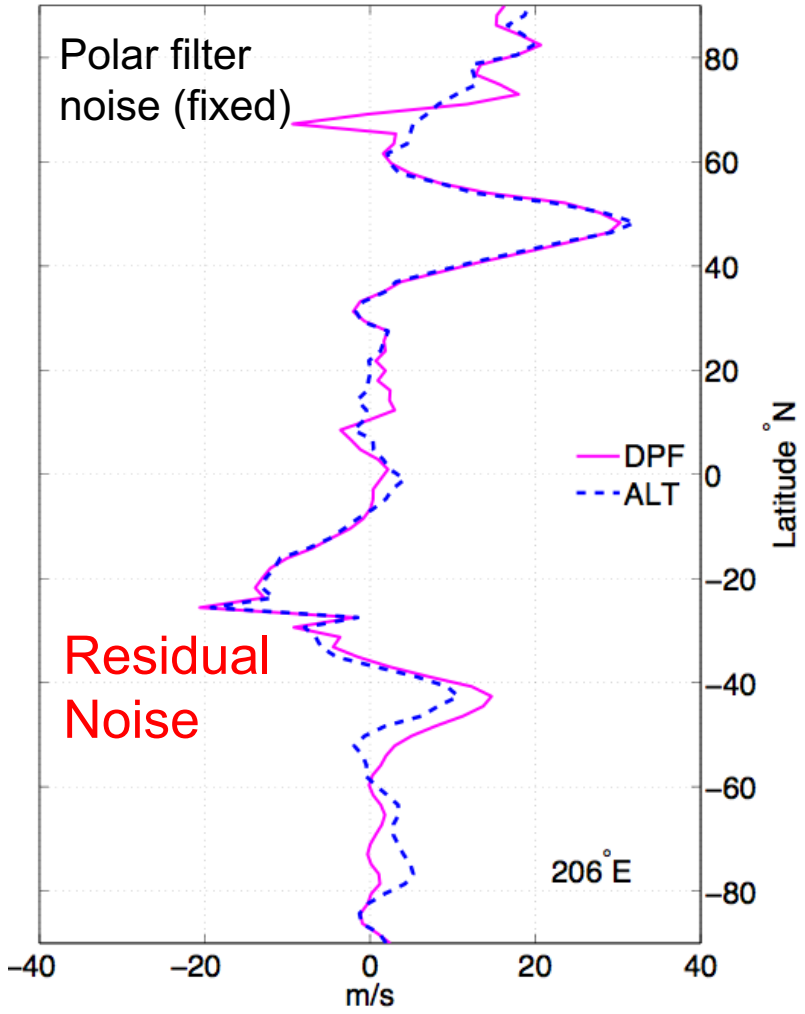
Ensemble Mean V @ 266hPa CAM FV core 00Z 25 September 2006

North-South cross sections

46º East



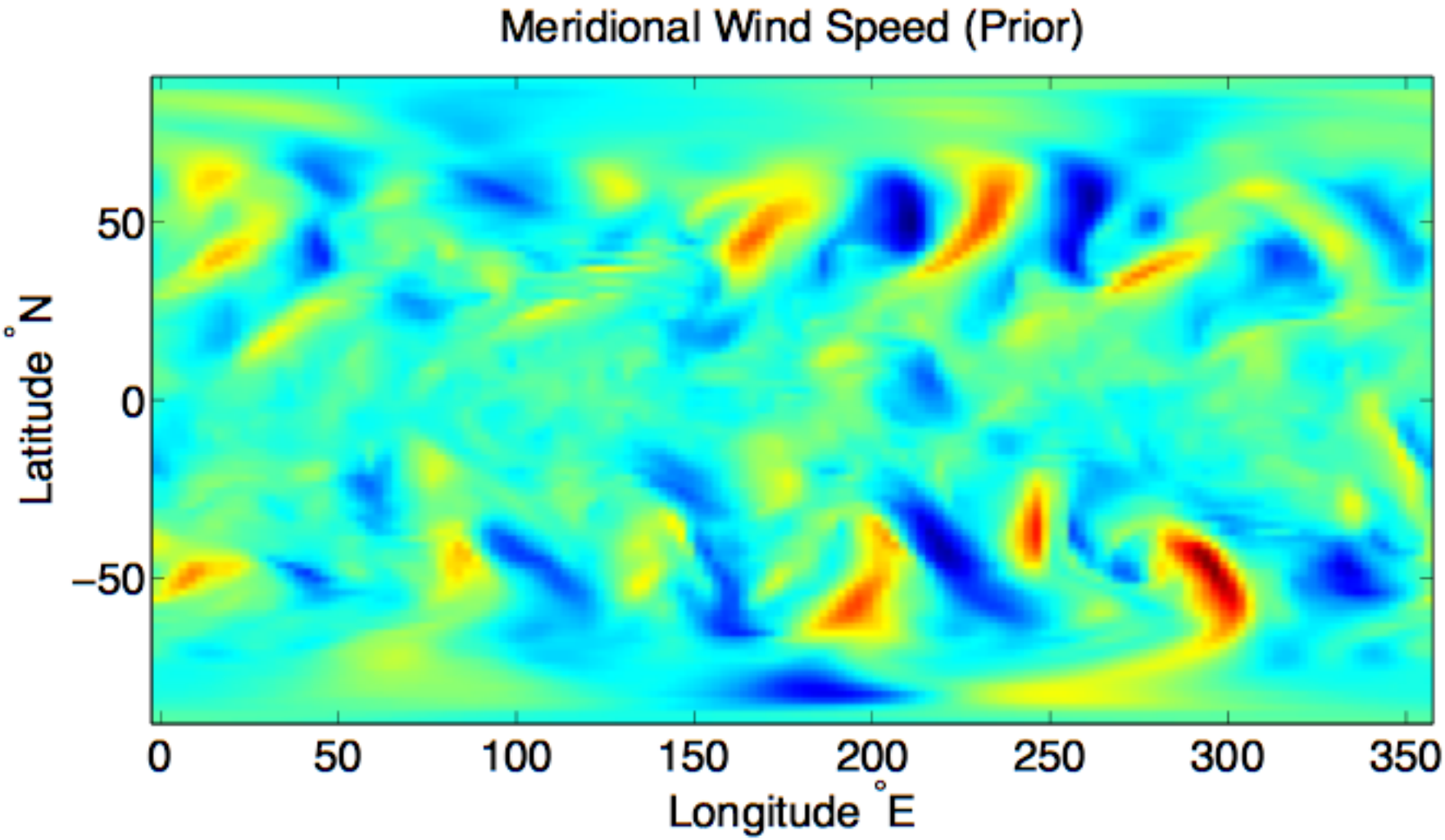
206º East



Ensemble Mean V @ 266hPa CAM FV core 00Z 25 September 2006

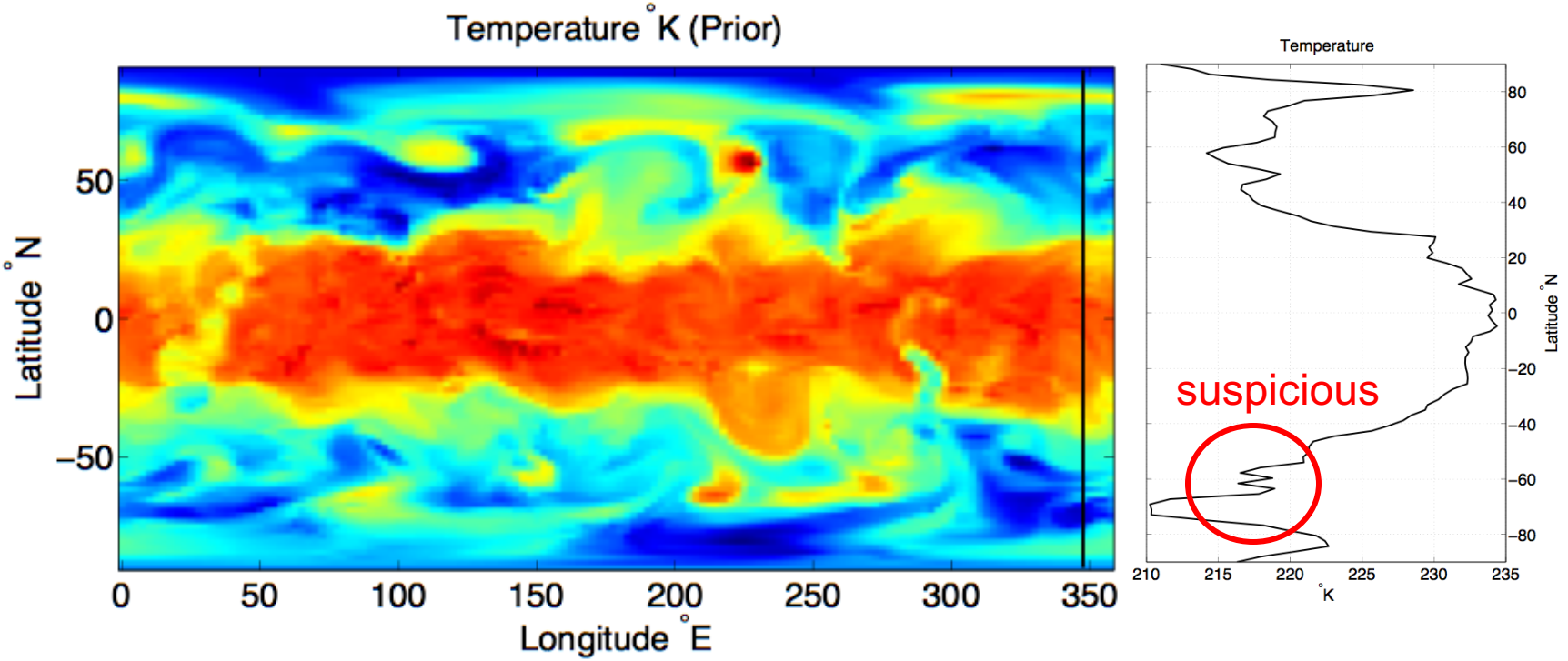
Another instance of noise from real-time use of DART-CAM in a chemistry field campaign (ARCTAS)

6 hour forecast of a single ensemble member

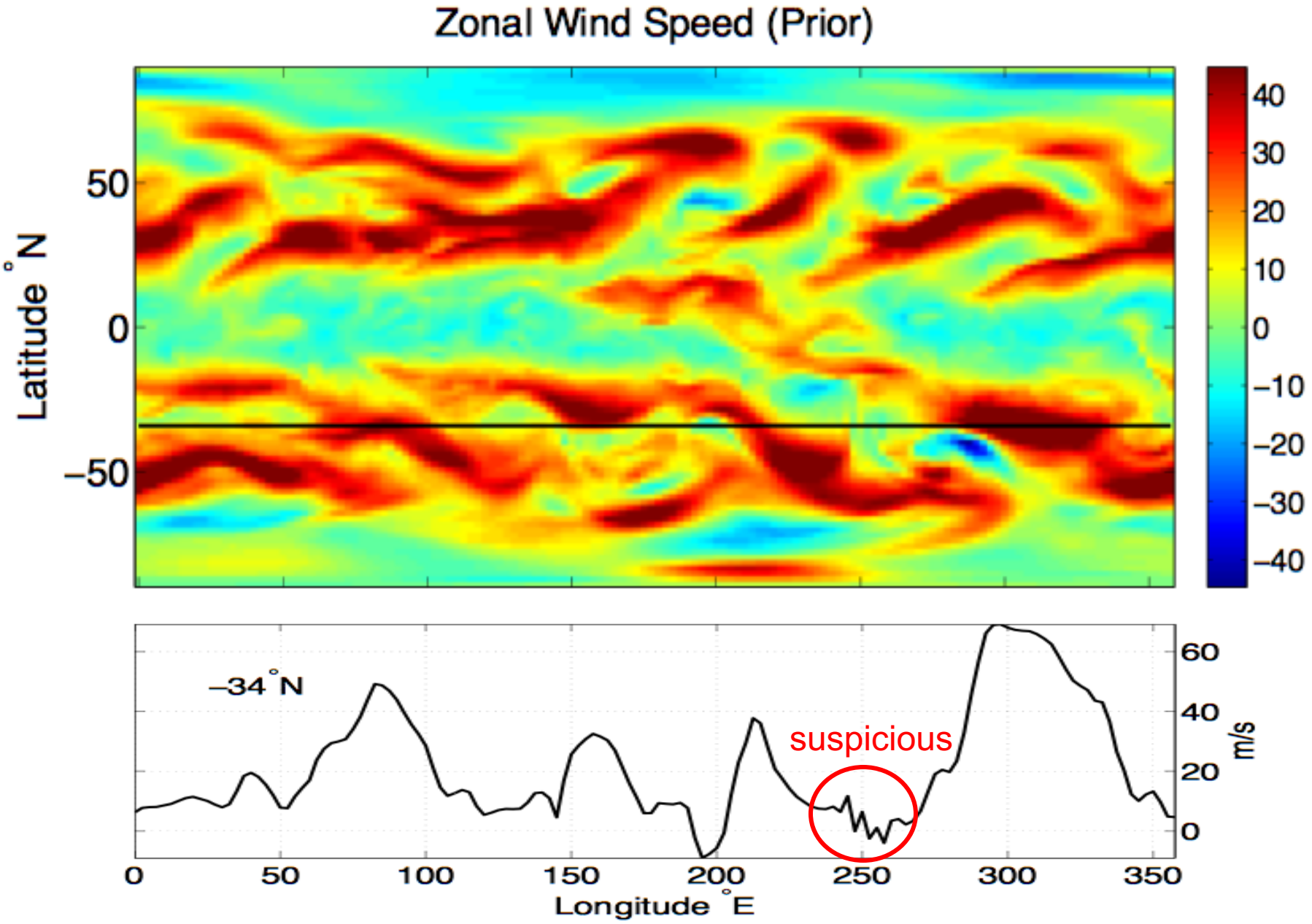


Ensemble Member 10 V @ 266hPa CAM FV core 06Z 13 April 2008

Noise not restricted to V winds …

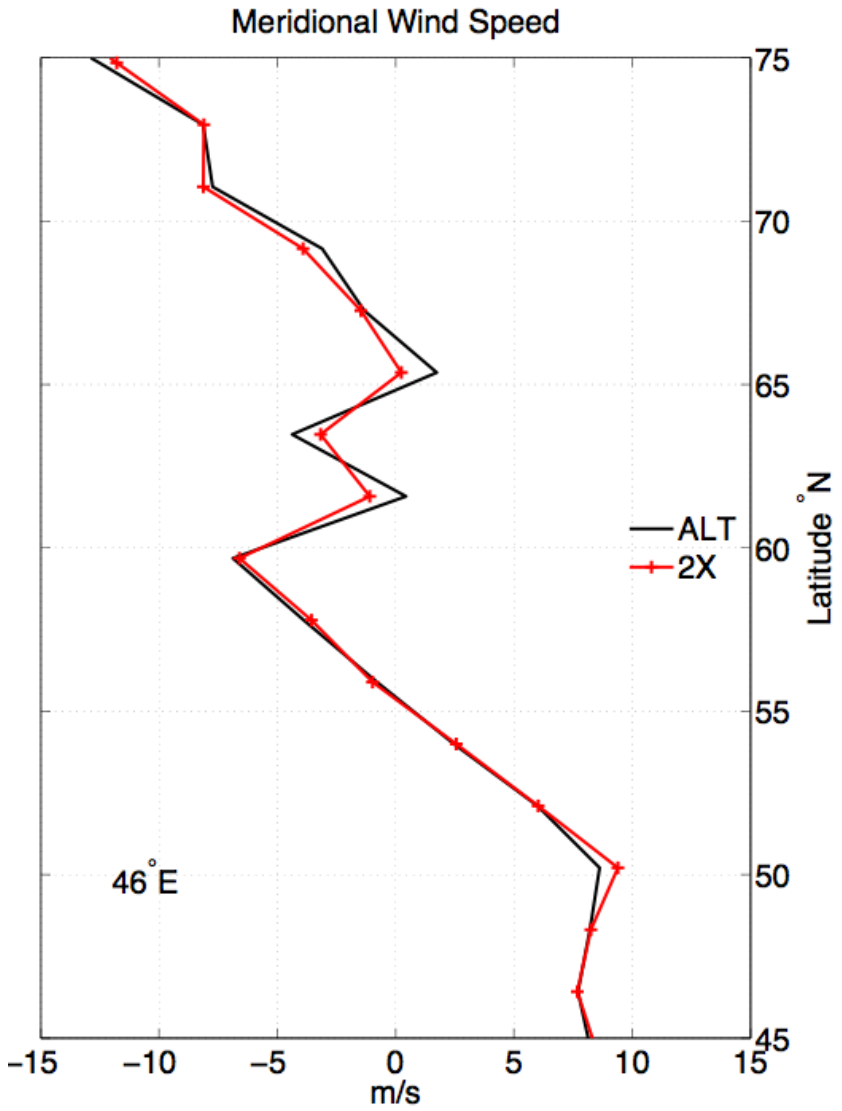


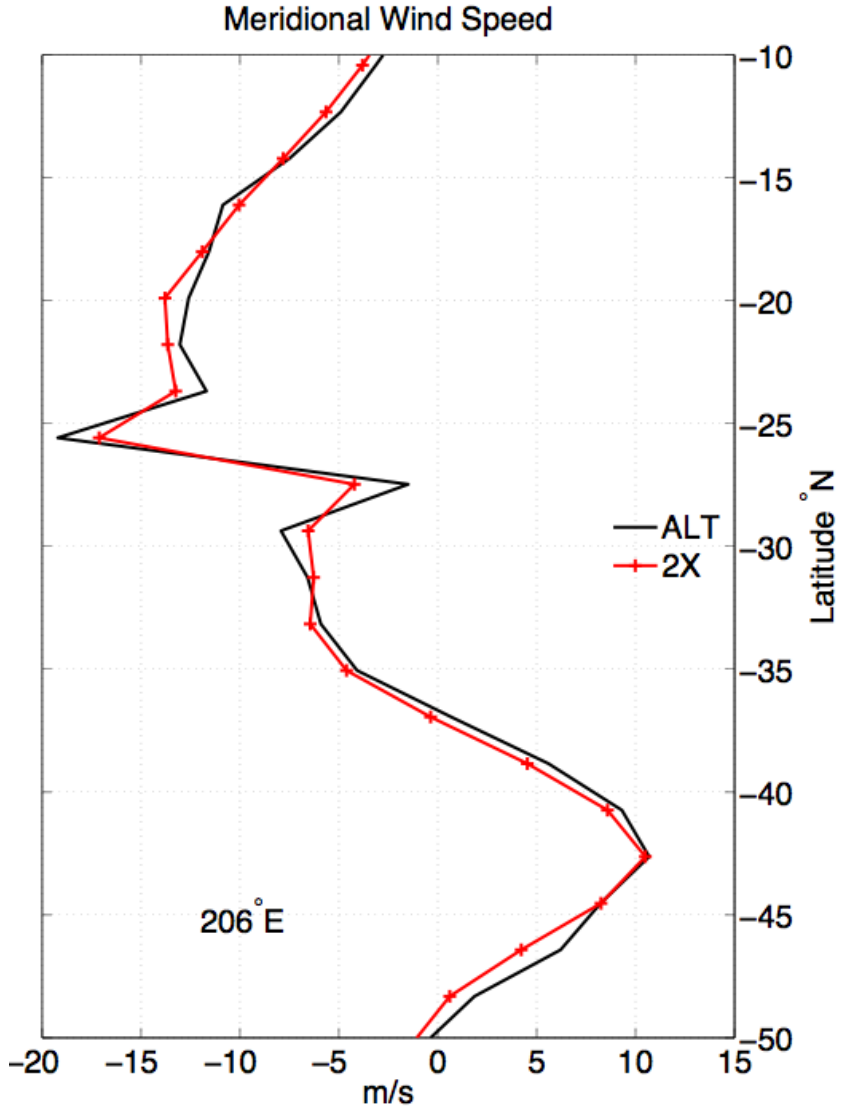
Ensemble Member 10 T @ 266hPa CAM FV core 06Z 13 April 2008



Ensemble Member 10 U @ 266hPa CAM FV core 06Z 13 April 2008

Doubling the dynamical time splitting reduced the noise; implicates model as opposed to assimilation.





Ensemble Mean V @ 266hPa CAM FV core 00Z 25 September 2006

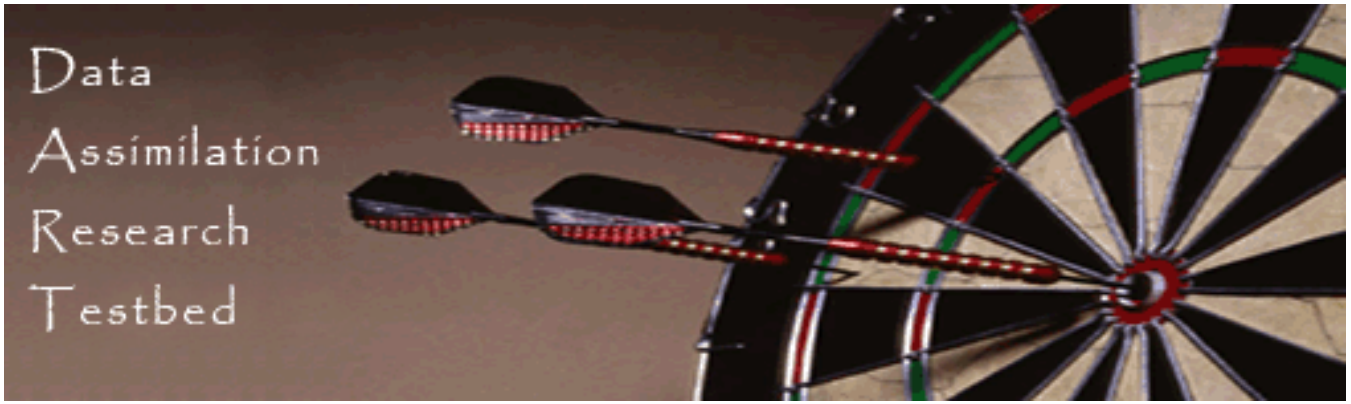
Notes and Conclusions

The noise here may seem small and transient,
but since it had not been recognized by any of the labs
that used this FV core, the effects on climate
runs were not explored.

- Spurious mixing is happening.
- Parameterizations may have been mistuned.
- More time may need to be spent fixing the remaining noise and looking at other unexamined pieces of the code.

Conclusions

- Existing GCMs contain unknown errors.
- These errors affect model results in unknown ways.
- Prediction using data assimilation can help reveal some model problems.
- Tracing problems to errors can be extremely difficult.
- Vastly improved development process is required for trustworthy models.



www.image.ucar.edu/DAReS/DART