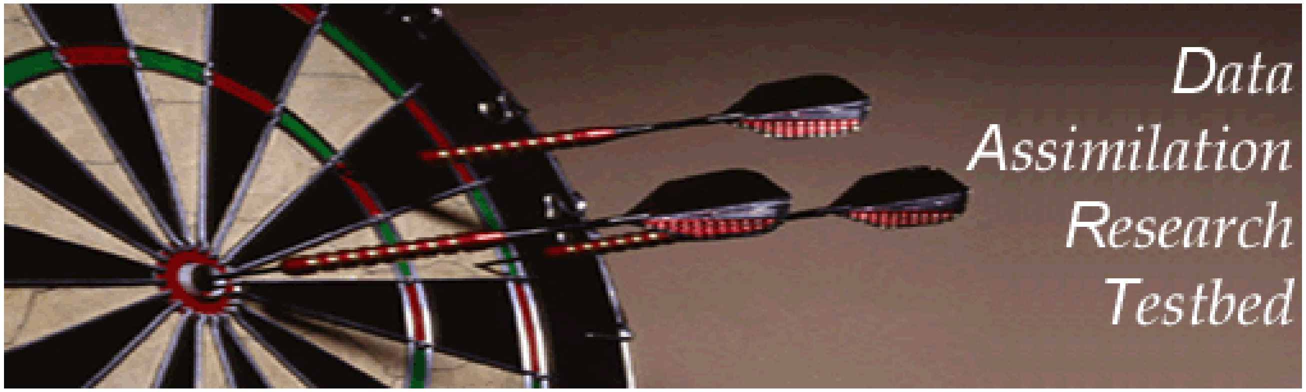
Data Assimilation Research Testbed Tutorial



Section 18: Lost in Phase Space:
The Challenge of Not Knowing the Truth

Version 1.0: June, 2005

In real applications, the truth is unknown.

All that we have is observations.

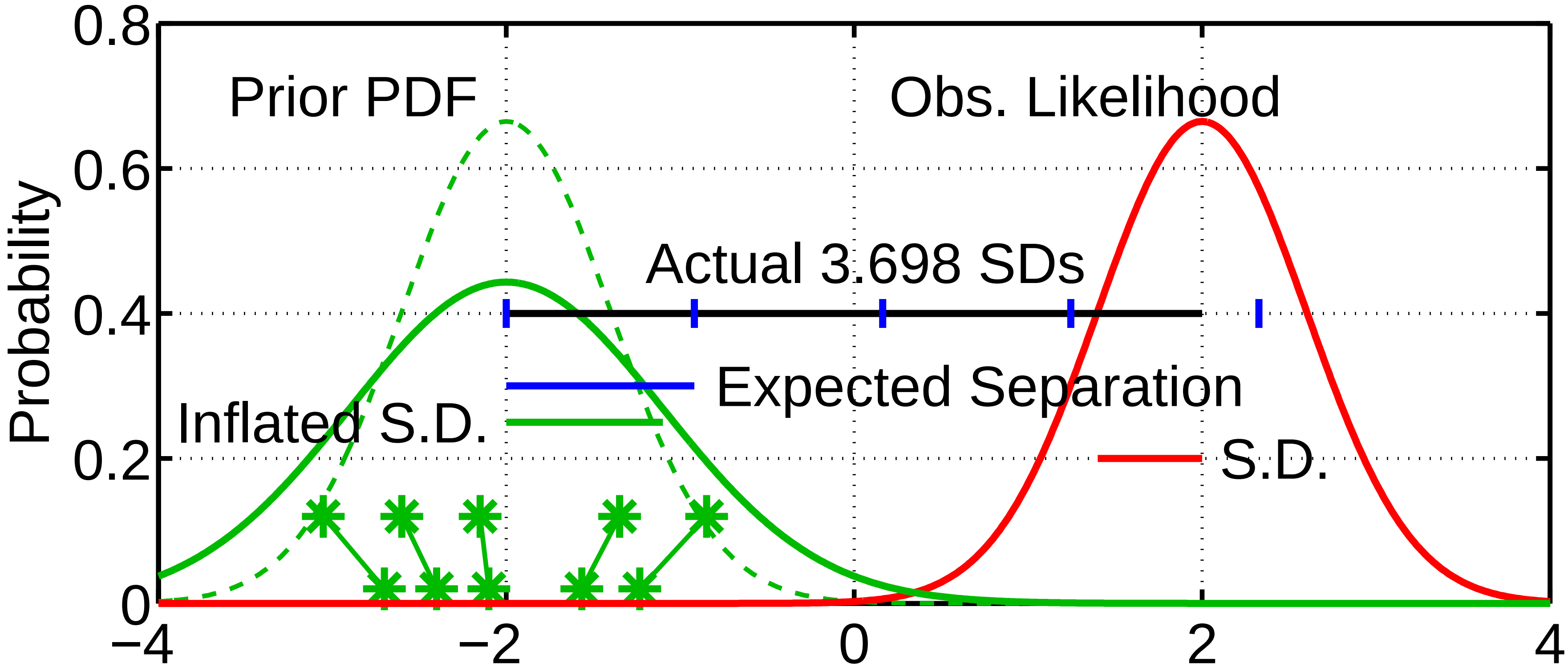
Having the truth available has been convenient, but also misleading...

Much less information is available from the observations.

They are generally functions of the state variables.

They are always contaminated with observational errors.

Recall that Expected(prior mean - observation) = $\sqrt{\sigma^2_{prior} + \sigma^2_{obs}}$.



Error is dominated by observational noise if $\sigma_{obs} >> \sigma_{prior}$.

Suppose σ_{obs}=1.0, σ_{prior} = 0.1 => E(RMS) = 1.005.

Halving σ_{prior} to 0.05 => E(RMS) = 1.001; only a 0.4% reduction!

Limited Observation space diagnostics in DART:

Try them out in lorenz_96 model.

After an assimilation, the final observation info is in *obs_seq.final*.

The program *obs_diag* post-processes this file.

The *obs_diag_nml* in *input.nml* controls this post-processing.

For low-order models, the only meaningful parameter is *rat_cri*.

This rejects 'bad' observations.

If $\left|\overline{y^p} - y^o\right| > rat\text{–}cri\sqrt{\sigma^2_{prior} + \sigma^2_{obs}}$, don't use this observation.

This is EXTREMELY DANGEROUS, but useful.

Rejecting 'good' observations can lead to inflated estimate of quality.

Displaying obs_space diagnostics with matlab:

Following commands are available for low-order models:

fit_ens_mean_time;	Mean obs. prior RMS error.
fit_ens_spread_time;	Mean obs. prior spread.
obs_num_time;	Number of observations used.
fit_mean_spread_time;	Mean error and spread on one plot.

Exercise with lorenz_96:

Pick a case that works relatively well and looks at obs space diags.
Pick a case that is similar, but clearly different with physical space diagnostics.
See if you can detect this difference with observation space diags.

Quick look at a real atmospheric application

Results from CAM Assimilation: January, 2003

Model:

CAM 3.0 T42L26

U,V, T, Q and PS state variables impacted by observations.

Land model (CLM 2.0) not impacted by observations.

Climatological SSTs.

Assimilation / Prediction Experiments:

80 member ensemble divided into 4 equal groups.

Initialized from a climatological distribution (huge spread).

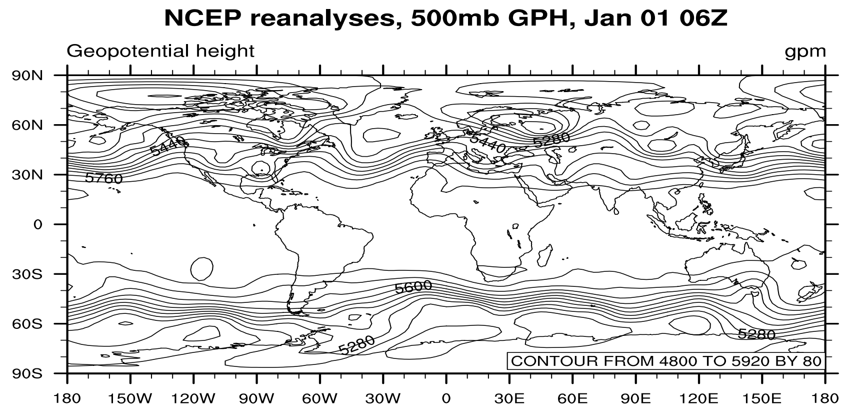
Initial tests for January, 2003.

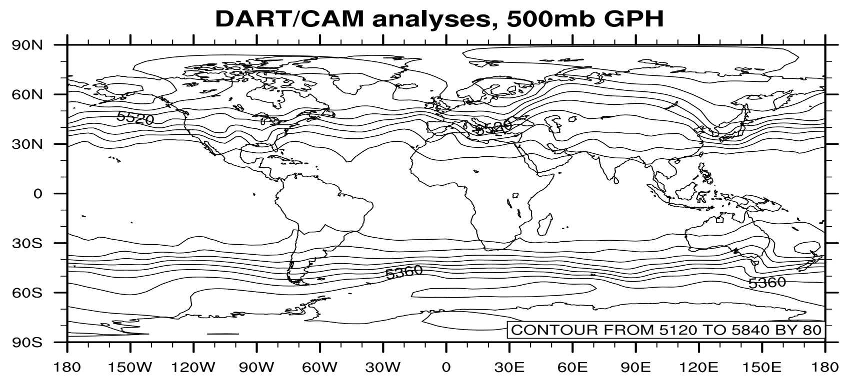
Uses most observations used in reanalysis

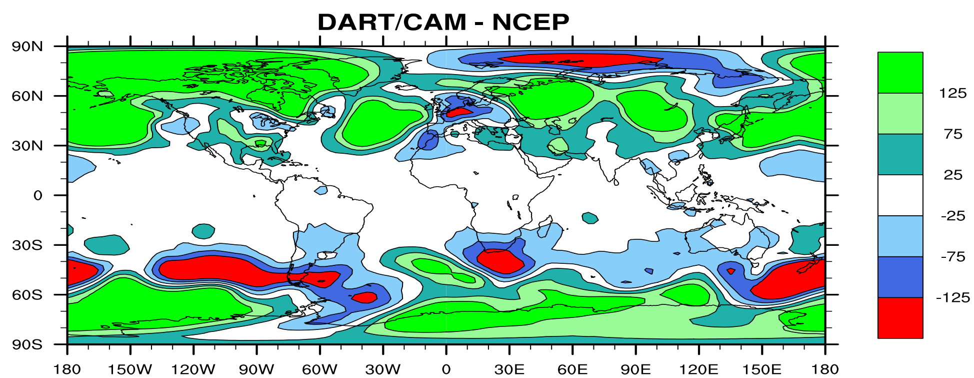
(Radiosondes, ACARS, Satellite Winds..., no surface obs.).

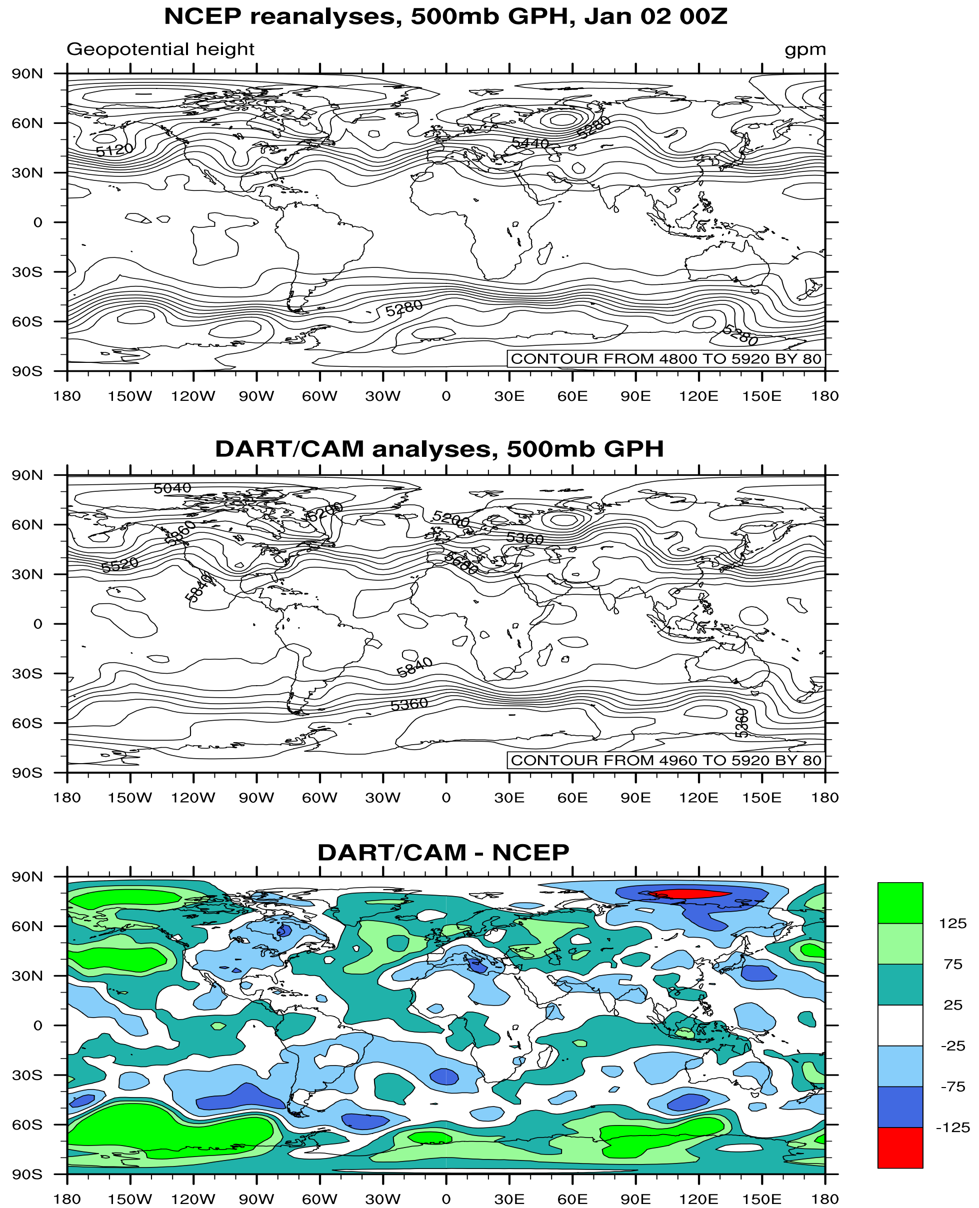
Assimilated every 6 hours; +/- 1.5 hour window for obs.

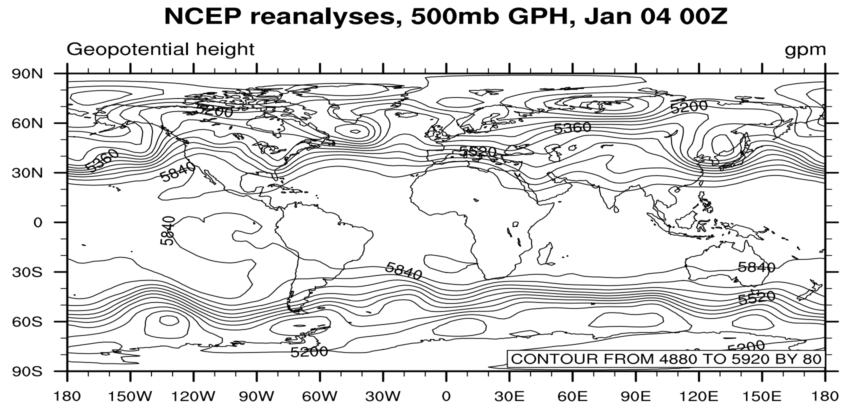
Adaptive error correction algorithm with fixed variance

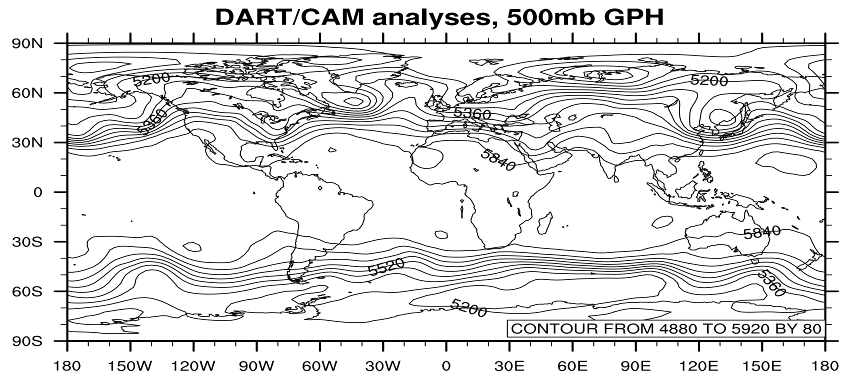

NCEP reanalyses, 500mb GPH, Jan 01 06Z
Geopotential height
gpm
90N
60N
30N
0
30S
60S
90S
180
150W
120W
90W
60W
30W
0
30E
60E
90E
120E
150E
180
5440
5760
5440
5280
5600
5280
5280
CONTOUR FROM 4800 TO 5920 BY 80

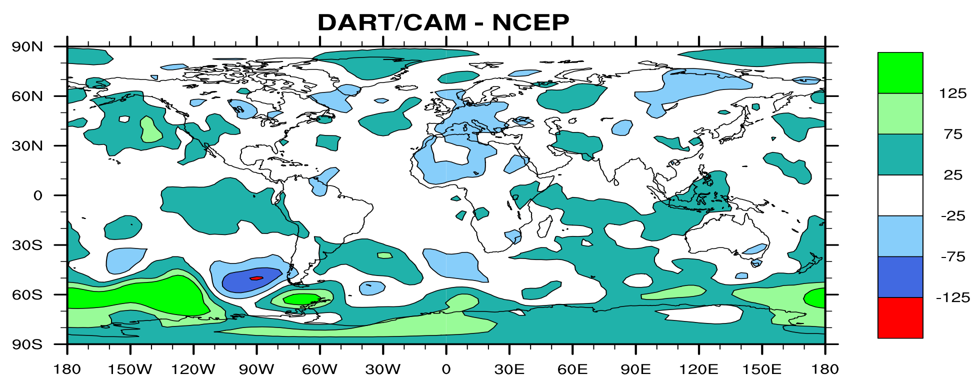

DART/CAM analyses, 500mb GPH
90N
60N
30N
0
30S
60S
90S
180
150W
120W
90W
60W
30W
0
30E
60E
90E
120E
150E
180
5520
5520
5360
5360
CONTOUR FROM 5120 TO 5840 BY 80

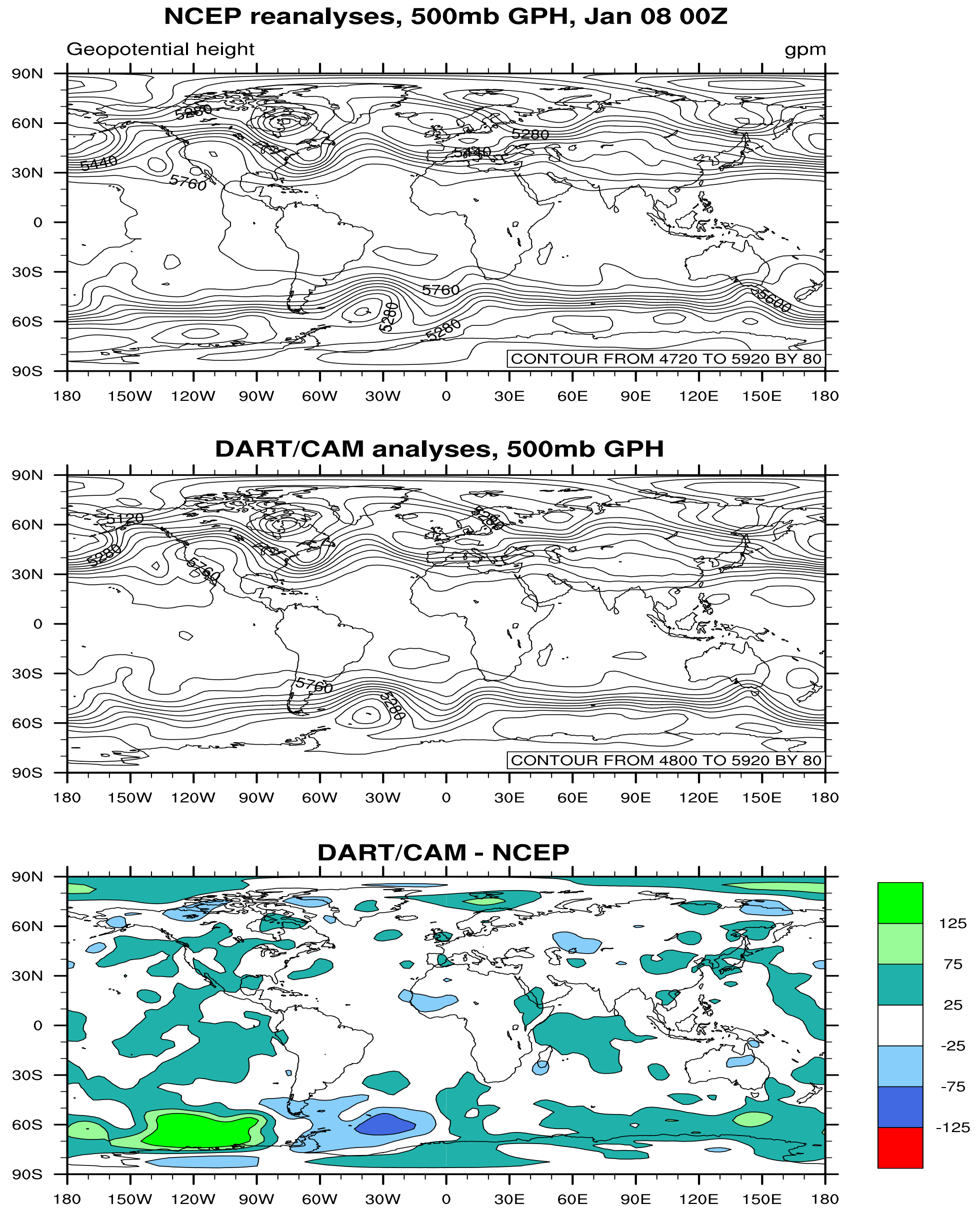

DART/CAM - NCEP
90N
60N
30N
0
30S
60S
90S
180
150W
120W
90W
60W
30W
0
30E
60E
90E
120E
150E
180
125
75
25
-25
-75
-125


NCEP reanalyses, 500mb GPH, Jan 02 00Z
Geopotential height
gpm
CONTOUR FROM 4800 TO 5920 BY 80
DART/CAM analyses, 500mb GPH
CONTOUR FROM 4960 TO 5920 BY 80
DART/CAM - NCEP
125
75
25
-25
-75
-125
90N
60N
30N
0
30S
60S
90S
180
150W
120W
90W
60W
30W
0
30E
60E
90E
120E
150E
180

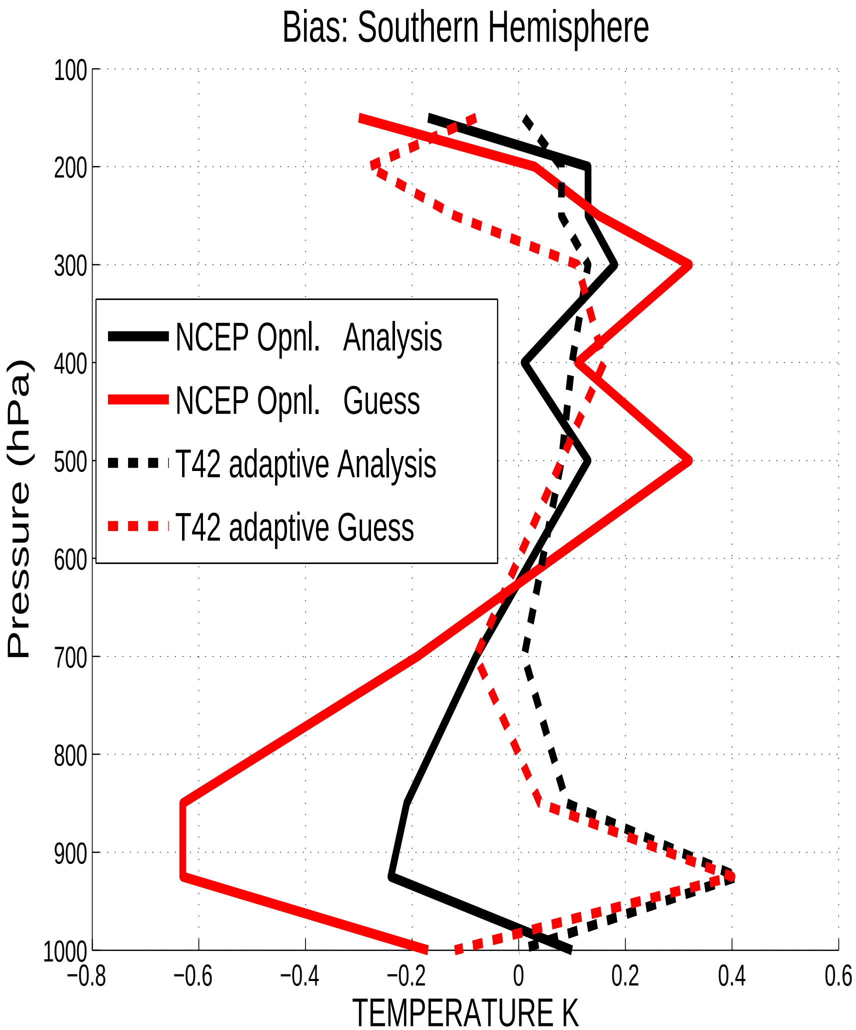

NCEP reanalyses, 500mb GPH, Jan 04 00Z
Geopotential height
gpm
90N
60N
30N
0
30S
60S
90S
180
150W
120W
90W
60W
30W
0
30E
60E
90E
120E
150E
180
5200
5360
5520
5840
CONTOUR FROM 4880 TO 5920 BY 80


DART/CAM analyses, 500mb GPH
90N
60N
30N
0
30S
60S
90S
180
150W
120W
90W
60W
30W
0
30E
60E
90E
120E
150E
180
5200
5360
5520
5840
CONTOUR FROM 4880 TO 5920 BY 80

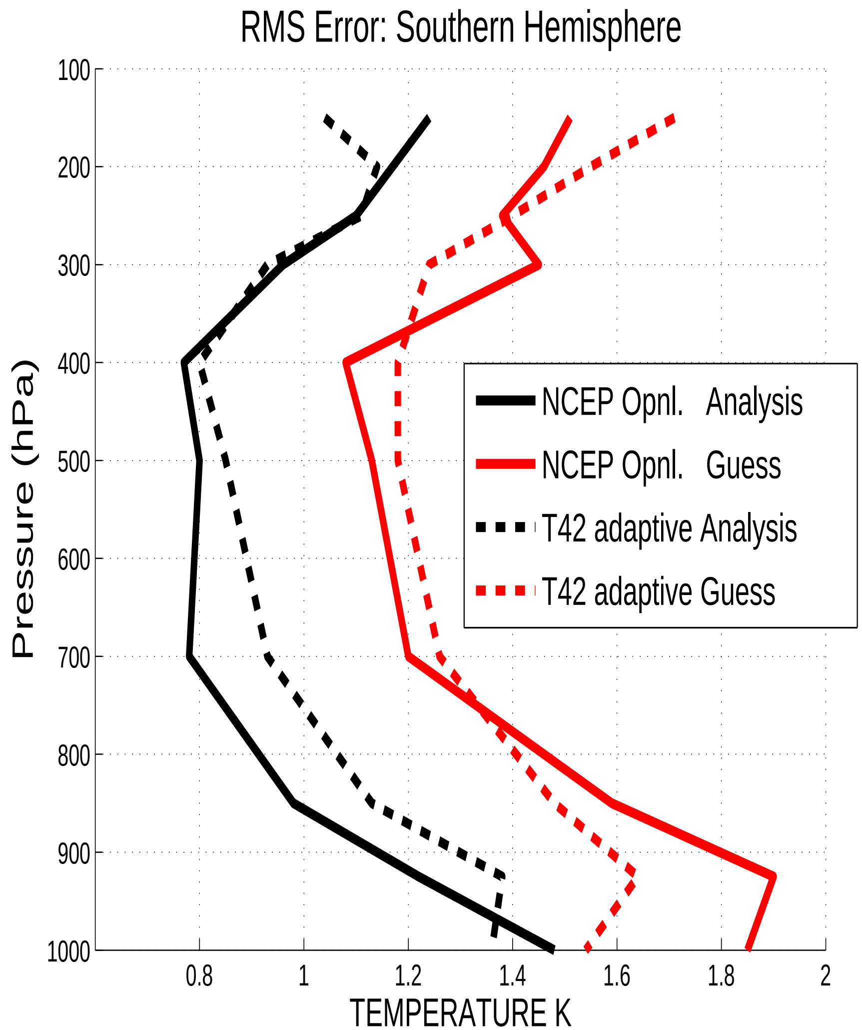

DART/CAM - NCEP
90N
60N
30N
0
30S
60S
90S
180
150W
120W
90W
60W
30W
0
30E
60E
90E
120E
150E
180
125
75
25
-25
-75
-125


NCEP reanalyses, 500mb GPH, Jan 08 00Z
Geopotential height
gpm
CONTOUR FROM 4720 TO 5920 BY 80
DART/CAM analyses, 500mb GPH
CONTOUR FROM 4800 TO 5920 BY 80
DART/CAM - NCEP
125
75
25
-25
-75
-125
90N
60N
30N
0
30S
60S
90S
180
150W
120W
90W
60W
30W
0
30E
60E
90E
120E
150E
180

Southern Hemisphere Temperature: Bias and RMSE

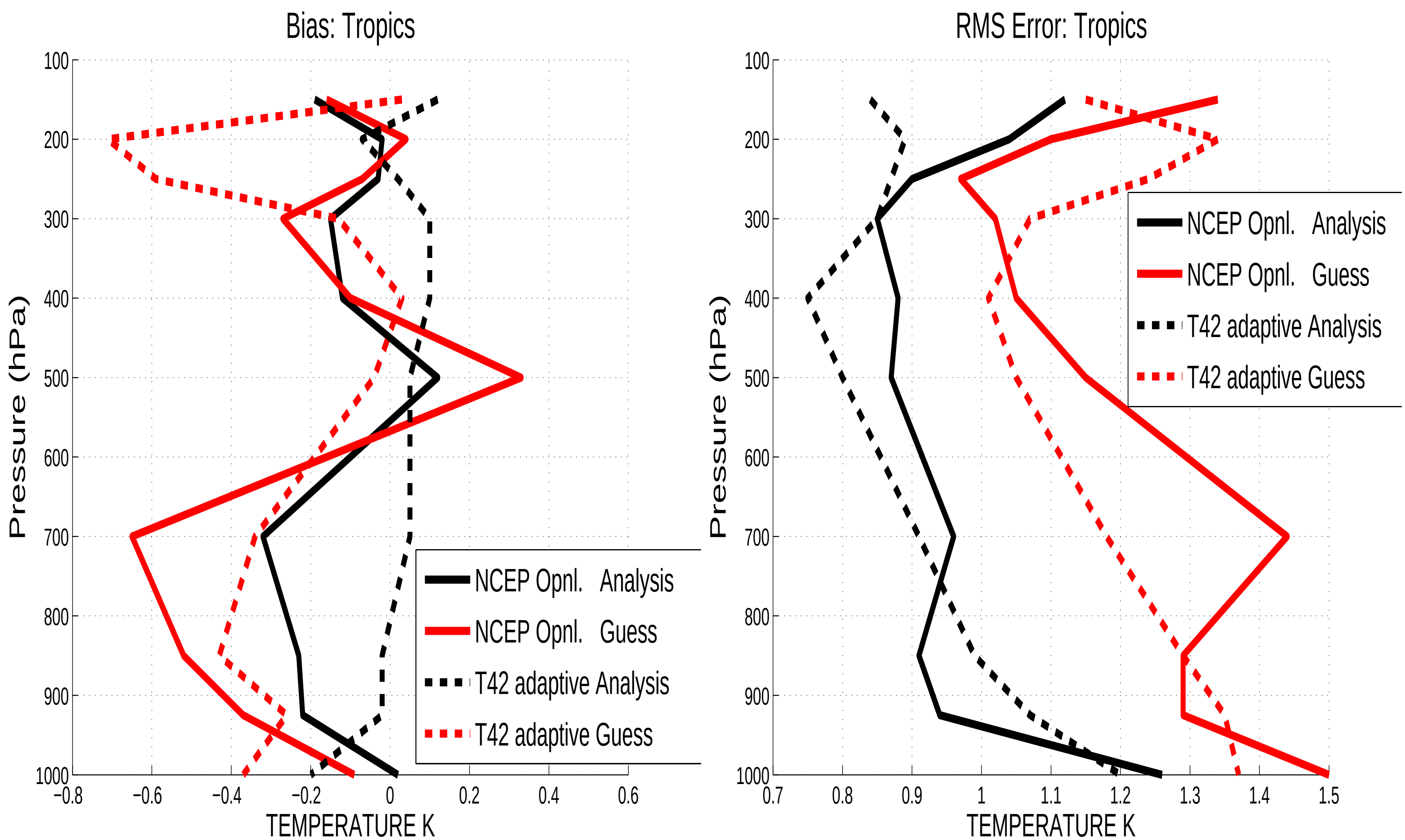


<u>Bias</u>



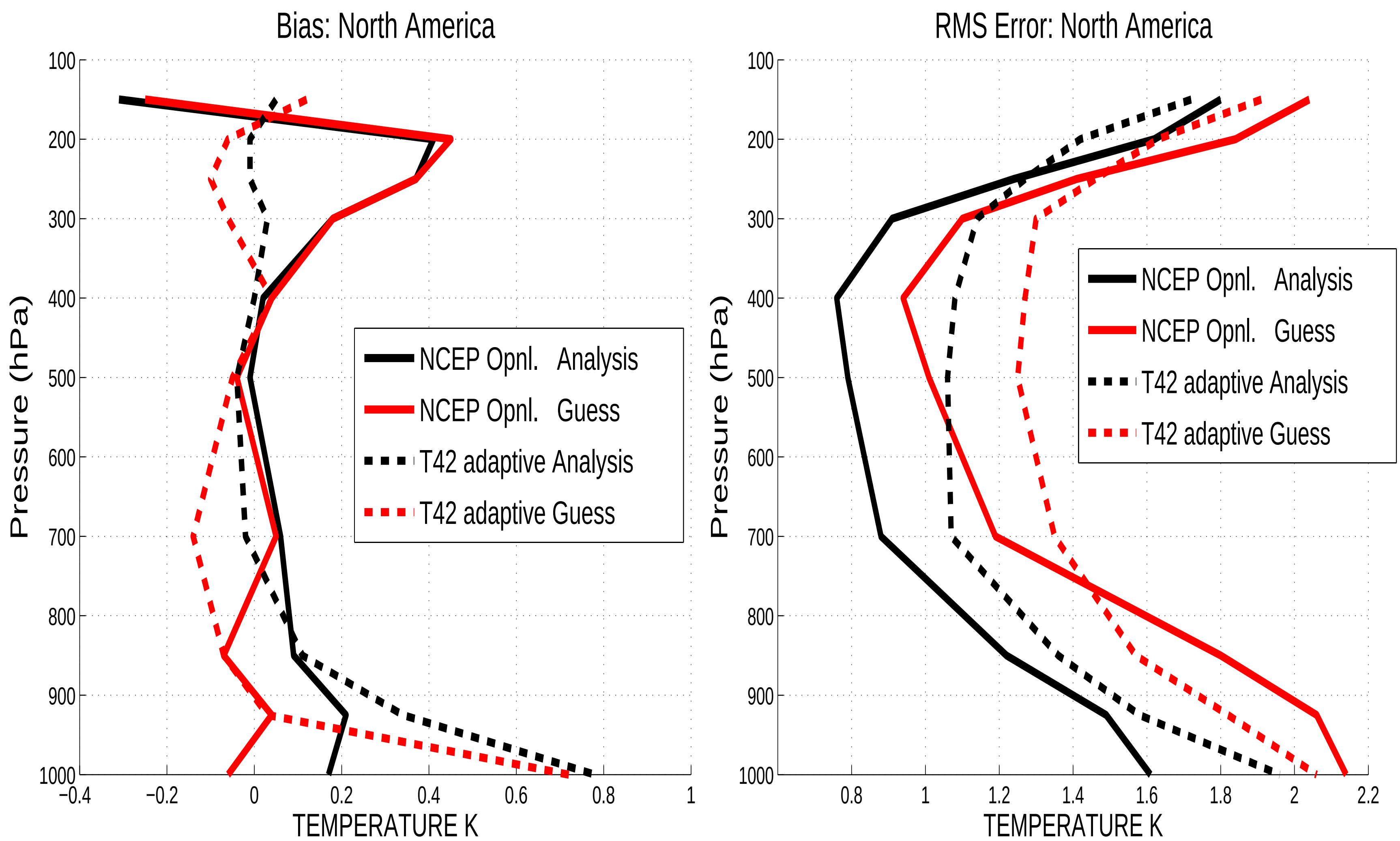
RMS

Tropics Temperature: Bias and RMSE



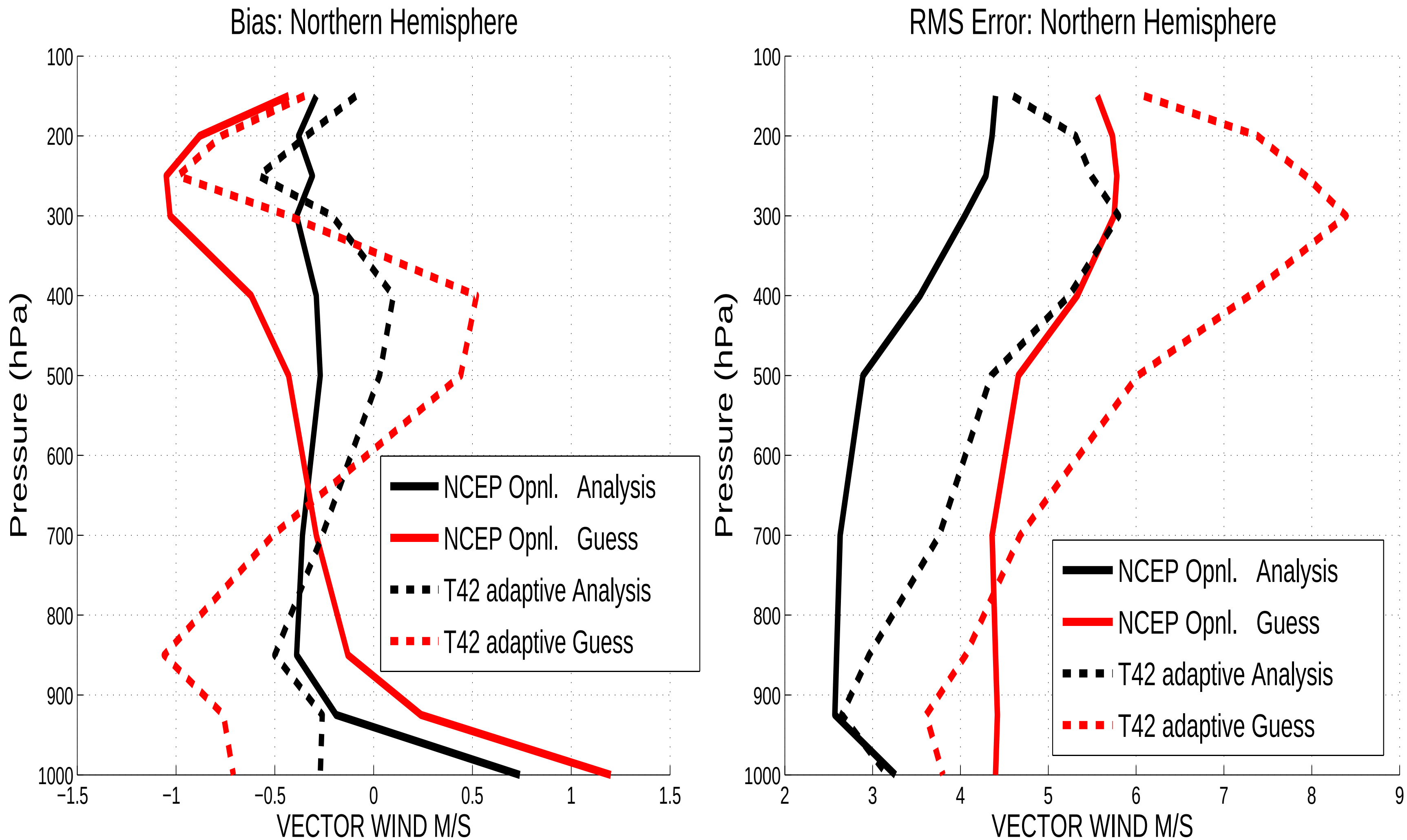
<u>Bias</u> **<u>RMS</u>**

North America Temperature: Bias and RMSE



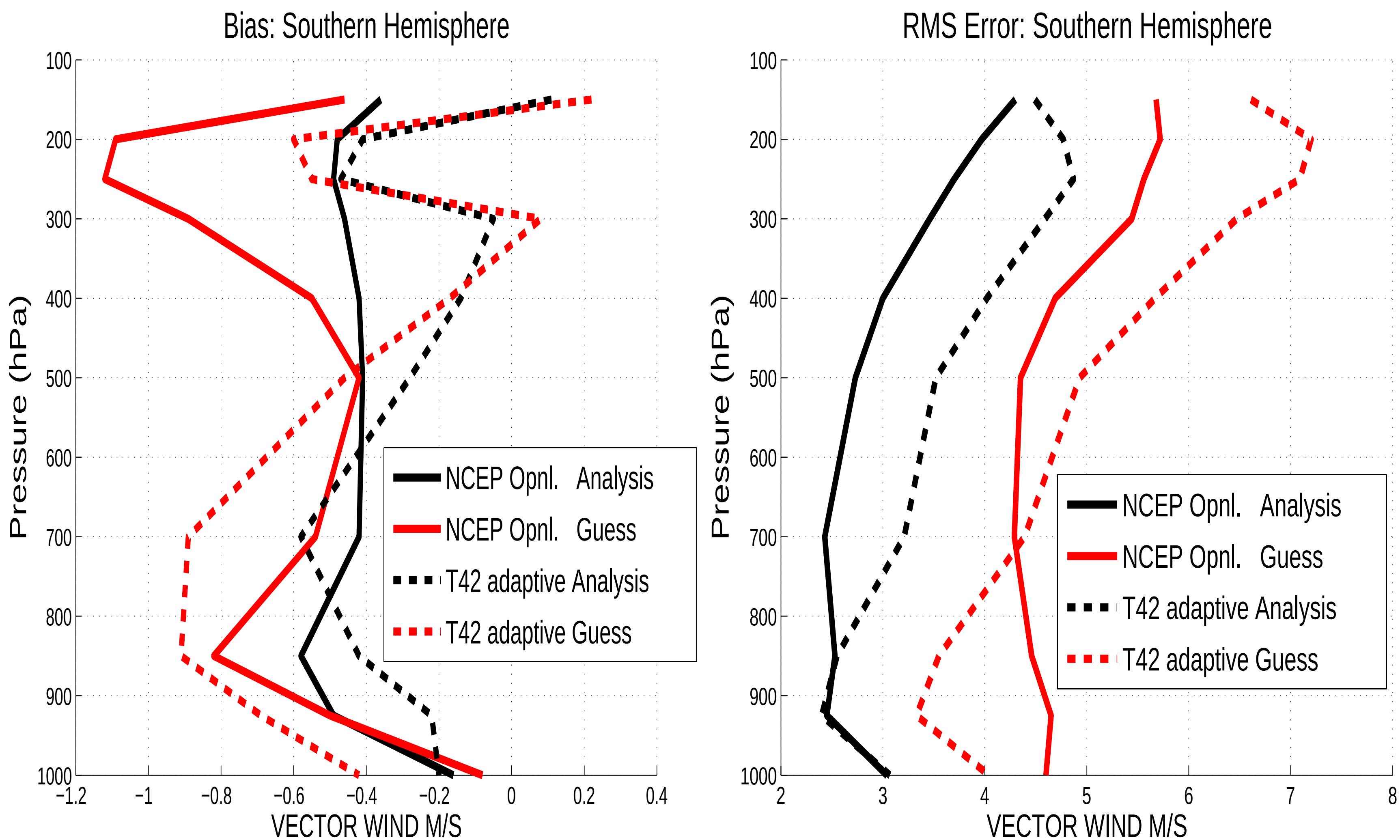
<u>Bias</u> <u>**RMS**</u>

Northern Hemisphere Wind: Bias and RMSE



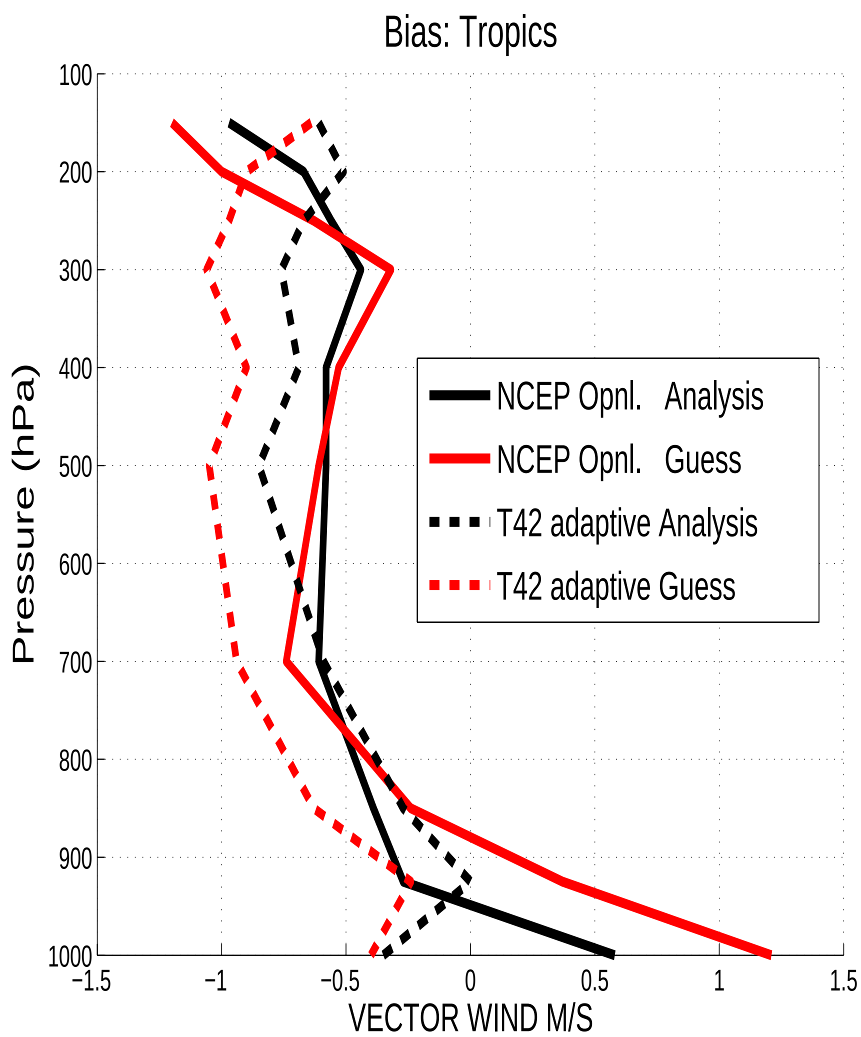
<u>Bias</u> **<u>RMS</u>**

Southern Hemisphere Wind: Bias and RMSE

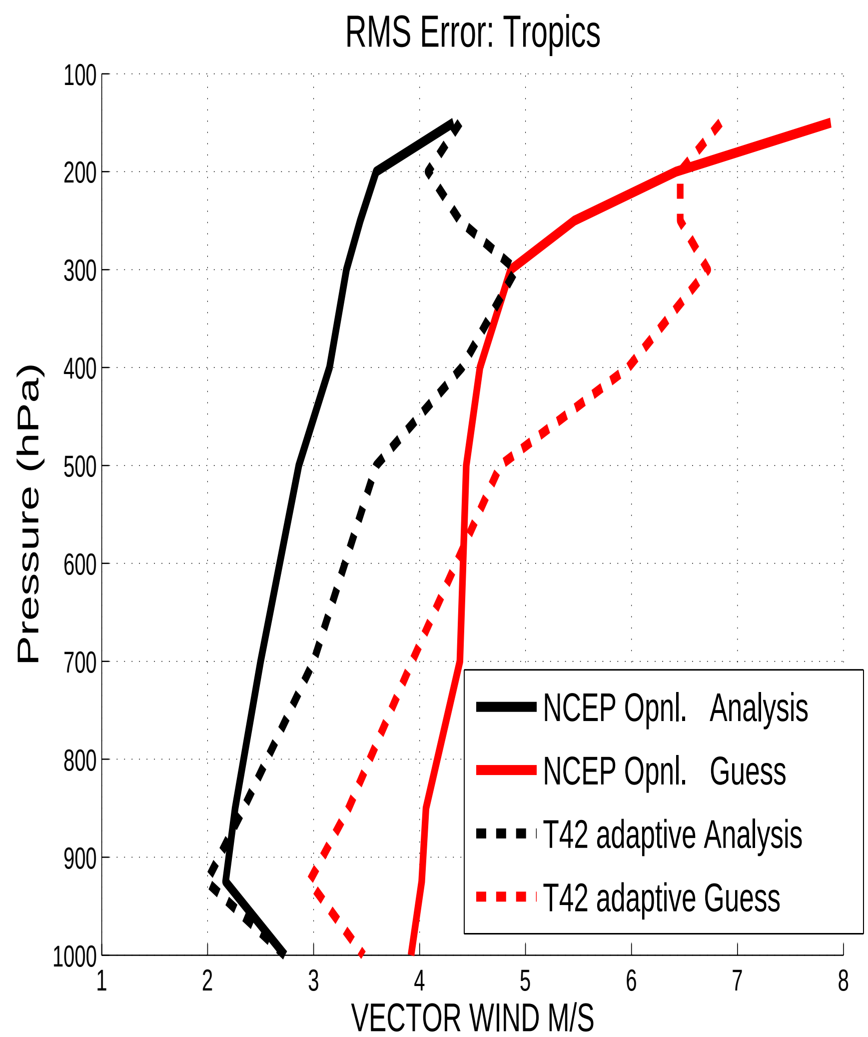


<u>Bias</u> **RMS**

Tropical Wind: Bias and RMSE

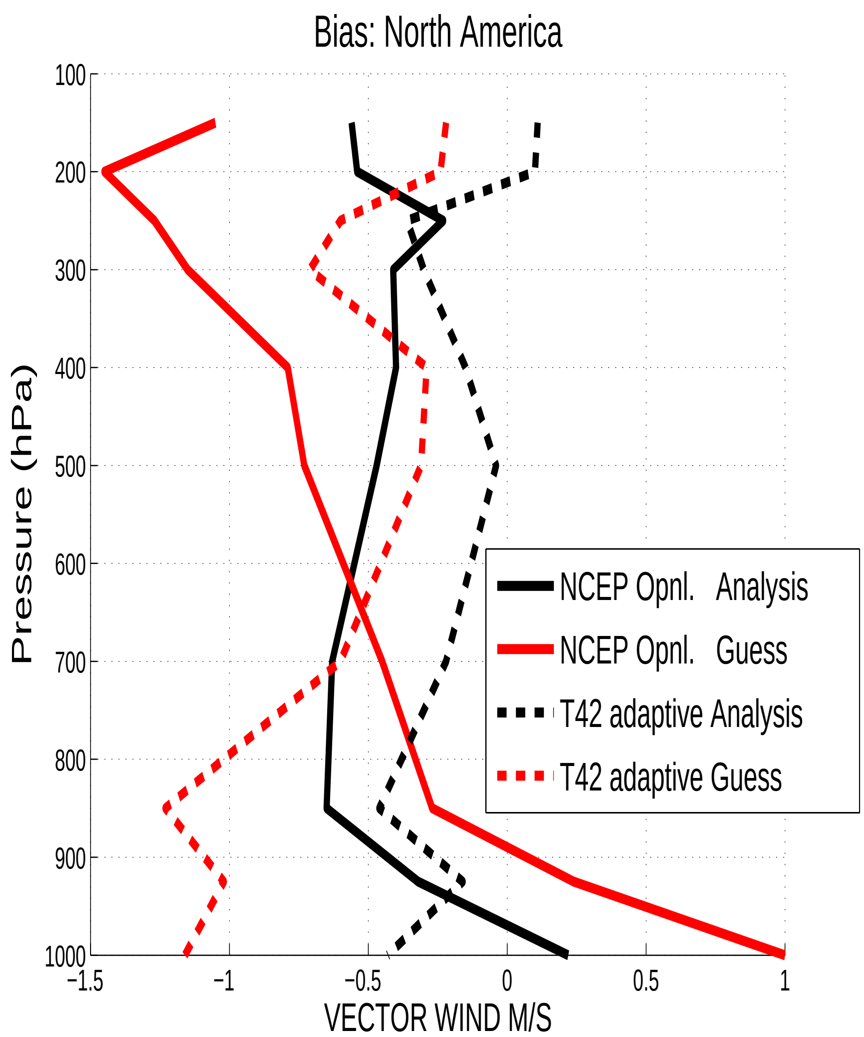


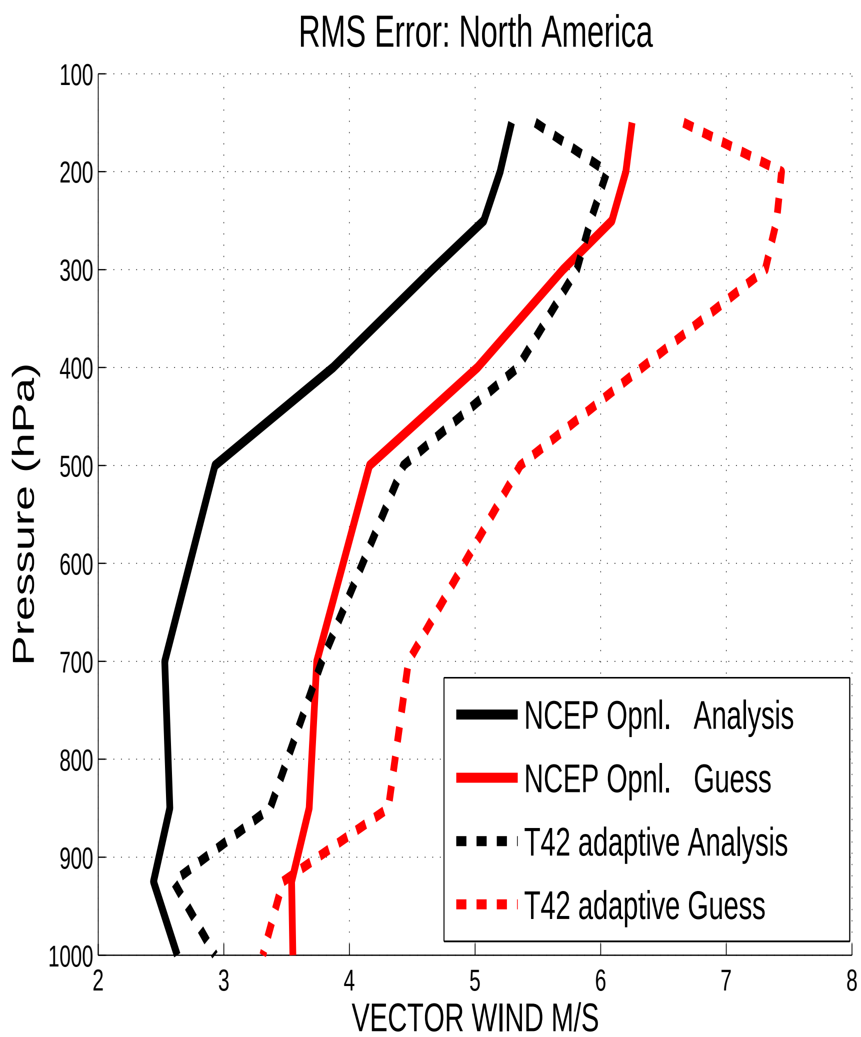
Bias



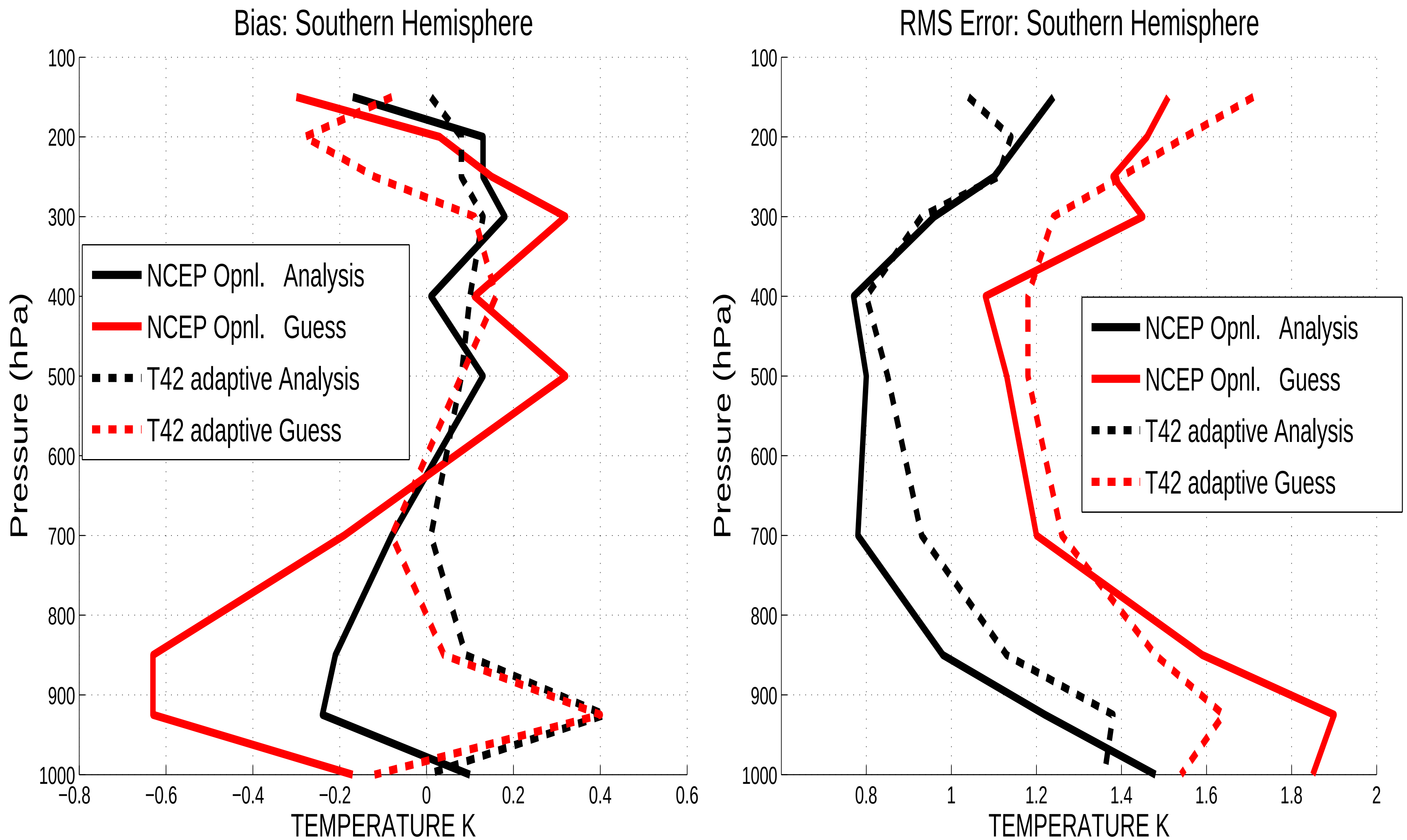
RMS

Northern America Wind: Bias and RMSE





Bias RMS


Bias: Southern Hemisphere
RMS Error: Southern Hemisphere
Pressure (hPa)
TEMPERATURE K
NCEP Opnl. Analysis
NCEP Opnl. Guess
T42 adaptive Analysis
T42 adaptive Guess
100
200
300
400
500
600
700
800
900
1000
−0.8
−0.6
−0.4
−0.2
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2

Sneak Preview of T85 CAM Results