Data Assimilation Experiments using Quality Controlled AIRS Version 5 Temperature Soundings

Joel Susskind

NASA GSFC

Laboratory for Atmospheres

Sounder Research Team

Joint Center for Data Assimilation Seminar

September 17, 2008

Infrared Atmospheric Sounders

Measure upwelling thermal radiance in a number of spectral channels

Most channels are in opaque spectral regions that do not see the surface

15 μm and 4.3 μm CO_2 bands, 6.7 μm H_2O band, 9.6 μm O_3 band, etc.

(667 cm^{-1}) (2350 cm^{-1}) (1600 cm^{-1}) (1040 cm^{-1})

Provide information about atmospheric temperature and constituent profiles

Channels are characterized by central frequency ν_i and band pass $\Delta\nu_i$

Fields of view (FOV's) are roughly 15 km at nadir

IR and microwave sounders are complementary and often fly together

AIRS was launched on Eos Aqua in May 2002

2360 channel grating detector array spectrometer 650 cm^{-1} - 2665 cm^{-1}

$\nu_i / \Delta\nu_i \approx 1200$ $\Delta\nu_i$ goes from 0.5 cm^{-1} - 2.2 cm^{-1}

Spatial resolution ≈ 13 km at nadir from 705 km orbit

AIRS was accompanied by AMSU A

IR and Microwave Observations are Very Complementary

IR Strengths

- Best vertical resolution (accuracy) of T(p) in mid-lower troposphere
- Water vapor profile information up to the tropopause
- Best information about surface skin temperature
- Trace gas profile information

IR Limitations

- Most channel observations are strongly affected by clouds

MW Strengths

- MW observations are not affected by most clouds
- MW observations help in accounting for effects of clouds on IR observations
- Microwave soundings of T(p), q(p) can be produced in overcast conditions

MW Limitations

- Channels sensitive to lower troposphere are highly affected by variable surface emissivity

Approaches to Account for Clouds in the Field of View (FOV)

1) Avoid clouds

a) Do soundings in areas "thought to be clear" (about 5% of the time)

b) Assimilate only channel radiances "thought to be unaffected by clouds"

All stratospheric sounding channel radiances

Tropospheric sounding channel radiances that do not see down to cloud tops

1b) is the approach used operationally by ECMWF, NCEP

2) Include cloud radiative transfer model in radiance calculation

Needs detailed knowledge of cloud microphysical and geometric properties within FOV

Potentially useful for single layer thin cirrus clouds

3) Attempt to determine cloud cleared radiances $\hat{R}_i$ from observations in adjacent fields of view

$\hat{R}_i$ represents radiances sounder "would see" if no clouds were in the FOV

K cloud formations requires K+1 FOV's to obtain $\hat{R}_i$

We use approach 3) to analyze AIRS/AMSU data

$\hat{R}_i$, as well as T(p), q(p), are products, each with their own error estimates

Overview of AIRS/AMSU Retrieval Methodology

Physically based retrieval system

Independent of GCM except for surface pressure - used to compute expected radiances

Uses cloud cleared radiances $\hat{R}_i$ to determine the solution

- $\hat{R}_i$ represents what AIRS would have seen in the absence of clouds
- Derivation of $\hat{R}_i$ is updated in different steps of the retrieval process

Cloud parameters and OLR are determined consistent with retrieved state and observed radiances

Cloud parameters are generated under all conditions

Successful AIRS/AMSU retrievals are produced in up to 90% cloud cover

All successful AIRS/AMSU retrievals have error estimates

Goddard DAAC had been analyzing AIRS/AMSU data using AIRS Version 5 algorithm

- Retrievals are near real time
- Analyzed data from September 2002 through the present

Significant Improvements in AIRS Version 5.0 Algorithm

Physical retrieval algorithm

Improved radiative transfer parameterization accounts for Non-Local Thermodynamic Equilibrium (non-LTE) effects

Allows for complete use of 4.3 μm CO_2 sounding channels

Following theoretical considerations:

Most 15 μm CO_2 channels are used only for cloud clearing

Gives clear column radiances $\hat{R}_i$ for all channels

$\hat{R}_i$ for 4.3 μm CO_2 channels are used to determine temperature profile T(p)

This allows for accurate soundings under more difficult cloud conditions

Error estimates

New methodology developed to provide accurate case by case error estimates

Uses residuals of the physical retrieval and cloud clearing steps

Error estimates used directly for quality control

Cloud Clearing with a Single Cloud Layer

If fields of view 1 and 2 are otherwise identical but have differing amounts of a single cloud type

$$R_{i,1} = (1-\alpha_1)R_{i,CLR} + \alpha_1 R_{i,CLD}$$

$$R_{i,2} = (1-\alpha_2)R_{i,CLR} + \alpha_2 R_{i,CLD}$$

then $R_{i,CLR} = R_{i,1} + \eta\left(R_{i,1} - R_{i,2}\right)$ where $\eta = \alpha_1 / (\alpha_2 - \alpha_1)$

If we have an estimate of $R^0_{i,CLR}$ for channel i computed from state X^0, can solve for η^0

$$\eta^0 = \frac{R^0_{i,CLR} - R_{i,1}}{R_{i,1} - R_{i,2}}$$

and

$$\hat{R}^0_j = R_{j,1} + \eta^0\left(R_{j,1} - R_{j,2}\right) \text{ for all channels j}$$

$\hat{R}_j^0$ is used to derive updated state X^1

Methodology to obtain η - Dual Frequency Cloud Clearing Principle

Start with initial state $T^0(p) = T(p) + \delta T(p)$ where $\delta T(p)$ is the error in $T^0(P)$ - say $\delta T(p) = 1K$

Compute $R^0_{i,CLR}$ from $T^0(P)$

$$\delta\eta_i = \eta_i^0 - \eta = \frac{R^0_{i,CLR} - R_{i,CLR}}{R_{i,1} - R_{i,2}} \approx \frac{\delta R_{i,CLR}}{R_i} \approx \frac{[dB(\nu,T)/dT] \text{ x } \delta T}{B(\nu,T)}$$

$$\left[\frac{dB(\nu,T)}{dT} / B(\nu,T)\right] \approx \nu / T^2$$

For the same temperature profile error, $\delta\eta_i$ is proportional to ν_i

$\delta\eta$ computed from channel i at 730 cm^{-1} is $\left(\frac{730}{2390}\right)$ smaller than from channel i at 2390 cm^{-1}

A - If you use 15 μm channels for cloud clearing, and 4.2 μm channels to retrieve $T^1(p)$

$$T^1(p) - T(p) \approx 0.3\left[T^0(p) - T(p)\right]$$

B - If you use 15 μm channels (or 4.2 μm channels) for both

$T^1(p) - T(p) \approx T^0(p) - T(p)$ nothing is gained

Therefore it is optimal to do A

AIRS Cloud Clearing Procedure

Uses radiances in 9 fields of view R_{ij} channel i, FOV j within AMSU A FOR

Allows for up to 8 cloud formations

$\bar{R}_i$ = average radiance over 9 FOV's in a set of i cloud clearing channels

$$\hat{R}_i^n = \bar{R}_i + \sum_{j=1}^{9} \eta_j^n \left(R_{i,j} - \bar{R}_i\right) = \bar{R}_j + \sum_{j=1}^{9} \eta_j^n \Delta R_{i,j}$$

9 values of η_j determine $\hat{R}_i$ for all channels (8 are linearly independent)

We compute expected values of $R_{i,CLR}^n$ from a surface and atmosphere state X^n to obtain η^n

State X^n derived using both IR and MW observations

$$\eta_j^n = \left(\Delta R' N^{-1} \Delta R\right)^{-1} \Delta R' N^{-1} \Delta R_{CLR}^n$$

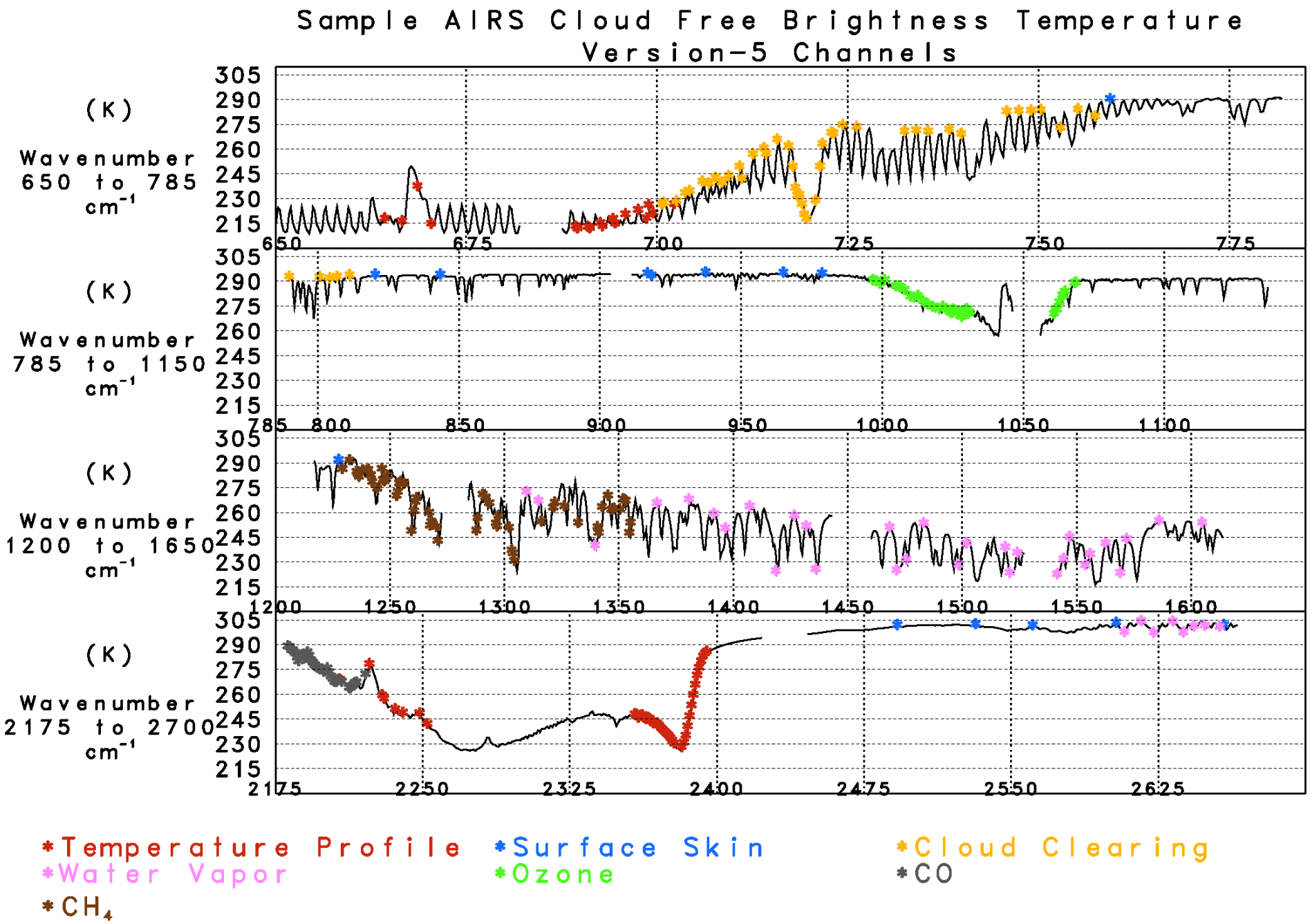
N = channel noise covariance matrix

where

$$\Delta R_{CLR,i}^n = R_{i,CLR}^n - \bar{R}$$

Use of $\hat{R}_i^n$ should in principle result in unbiased state X^{n+1} if X^n is unbiased

Using 15 μm channels to determine η_j^n and 4 μm channels to retrieve $T^{n+1}(p)$ minimizes effect of a bias in $T^n(p)$ on a resultant bias in $T^{n+1}(p)$


Sample AIRS Cloud Free Brightness Temperature
Version-5 Channels
(K)
Wavenumber 650 to 785 cm-1
Wavenumber 785 to 1150 cm-1
Wavenumber 1200 to 1650 cm-1
Wavenumber 2175 to 2700 cm-1
305
290
275
260
245
230
215
650
675
700
725
750
775
785
800
850
900
950
1000
1050
1100
1200
1250
1300
1350
1400
1450
1500
1550
1600
2175
2250
2325
2400
2475
2550
2625
Temperature Profile
Surface Skin
Cloud Clearing
Water Vapor
Ozone
CO
CH4

Objectives of AIRS/AMSU

Provide real time observations to improve numerical weather prediction

Could be R_i (used by NCEP, ECMWF) or T(p), q(p)

Accuracy of $\hat{R}_i$, T(p), q(p) degrades slowly with increasing cloud fraction

There is a trade-off between accuracy and spatial coverage

Using soundings or radiances only in clear cases limits utility of the data

Provide observations to measure and explain interannual variability and trends

Must provide good spatial coverage but also be unbiased

Can be less accurate than needed for data assimilation

Must not contain systematic data gaps in certain regions

AIRS Version 5 contains accurate error estimates for $\hat{R}_i$, T(p), and q(p)

Error estimates and quality flags provide options for use in either weather or climate applications

Generation of Empirical Error Estimates δX_i

This step is done after physical retrieval is otherwise completed

Methodology used for $\delta SST, \delta T(p), \delta W_{tot} / W_{tot}$ **is identical**

Uses 16 internally computed values of convergence tests Y_j

Thresholds of 12 Y_j terms were used in Version 4 quality control

δX_i, error estimate for X_i, is computed according to

$$\delta X_i = \sum M_{ij} Y_j$$

Determination of M_{ij}

Use profiles with "truth"

$$\Delta X_i = \left| X_i - X_i^{TRUTH} \right|$$

Each profile now has $\Delta X_i, Y_j$

M_{ij} found which minimizes RMS $\left| \delta X_i - \Delta X_i \right|$

M_{ij} generated using all September 29, 2004 cases in which IR retrieval is accepted

ECMWF taken as "truth" to provide ΔX_i

M_{ij} tested on January 25, 2003 - used once and for all

Same basic approach is used for $\delta \hat{R}_i, \delta q(p)$

Methodology Used for V5 Quality Control

Temperature Profile T(p)

Define a profile dependent pressure, p_g, above which the temperature profile is flagged as good - otherwise flagged as bad

Use error estimate $\delta T(p)$ to determine p_g

Start from 70 mb and set p_g to be the pressure at the first level below which $\delta T(p) >$ threshold $\Delta T(p)$ for 3 consecutive layers

Temperature profile statistics include errors of T(p) down to $p = p_g$

Sea surface temperature SST

Flag SST as good if $\delta SST < 1.0K$

Total precipitable water W_{tot}

Flag W_{tot} as good if $\delta W_{tot} / W_{tot} < 0.35$

Clear column radiance $\hat{R}_i$

Flag $\hat{R}_i$ as good if $\delta \hat{R}_i < 0.9K$ in brightness temperature error units

Thresholds for T(p) - Computation of p_g

p_g is the highest pressure at which $\delta T(p) > \Delta T(p)$ for 3 consecutive levels

$\Delta T(p)$ is defined at 3 pressures: $\Delta T(70\,\text{mb}), \Delta T(p_{surf/2})$, and $\Delta T(p_{surf})$

$\Delta T(p)$ is linearly interpolated in ln p between these 3 values

Separate threshold values for $\Delta T(p)$ are set for non-frozen ocean and for land/ice

Version 5 uses Standard thresholds optimized for weather and climate simultaneously

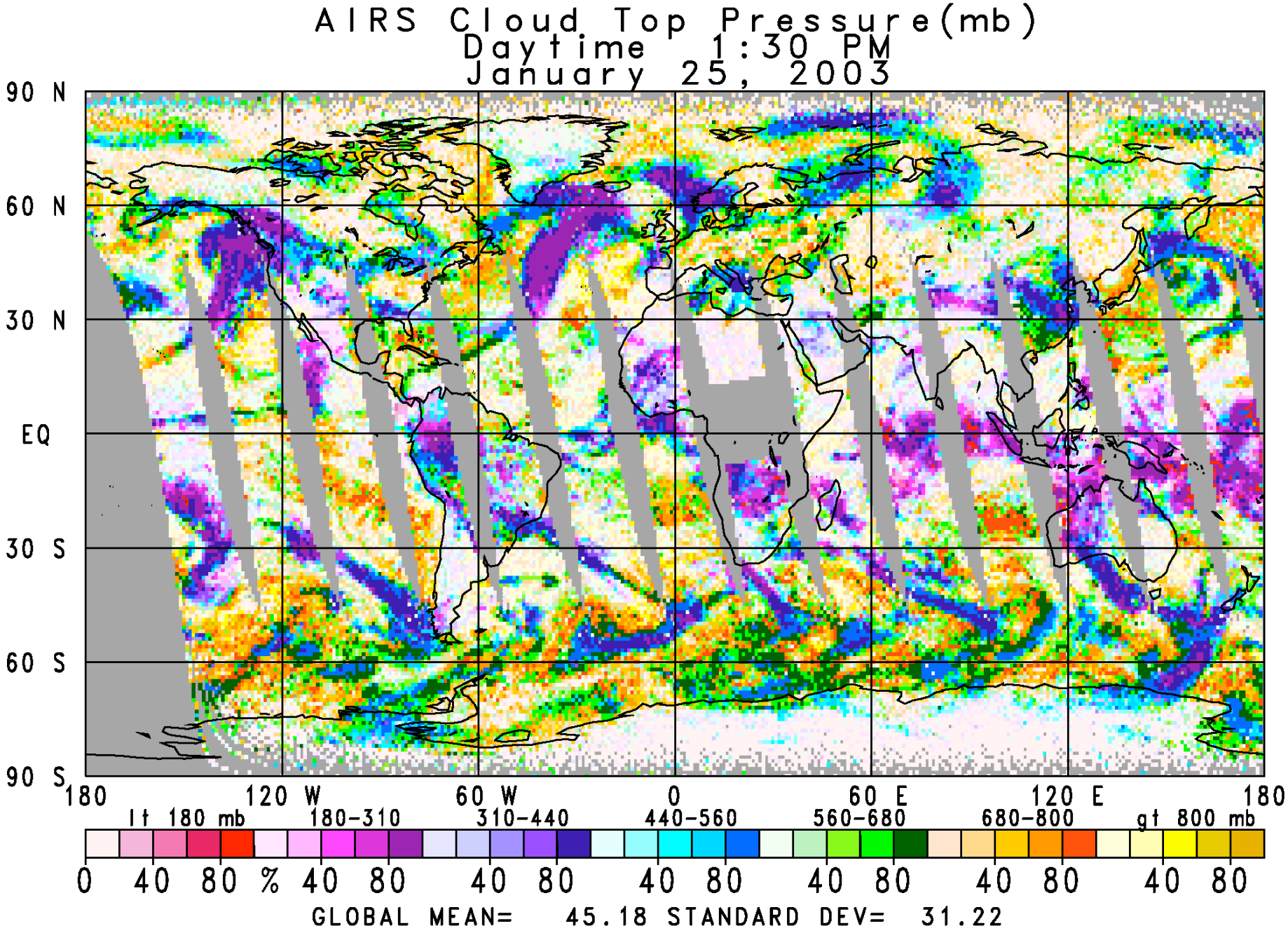
Goal was to produce good spatial coverage for climate

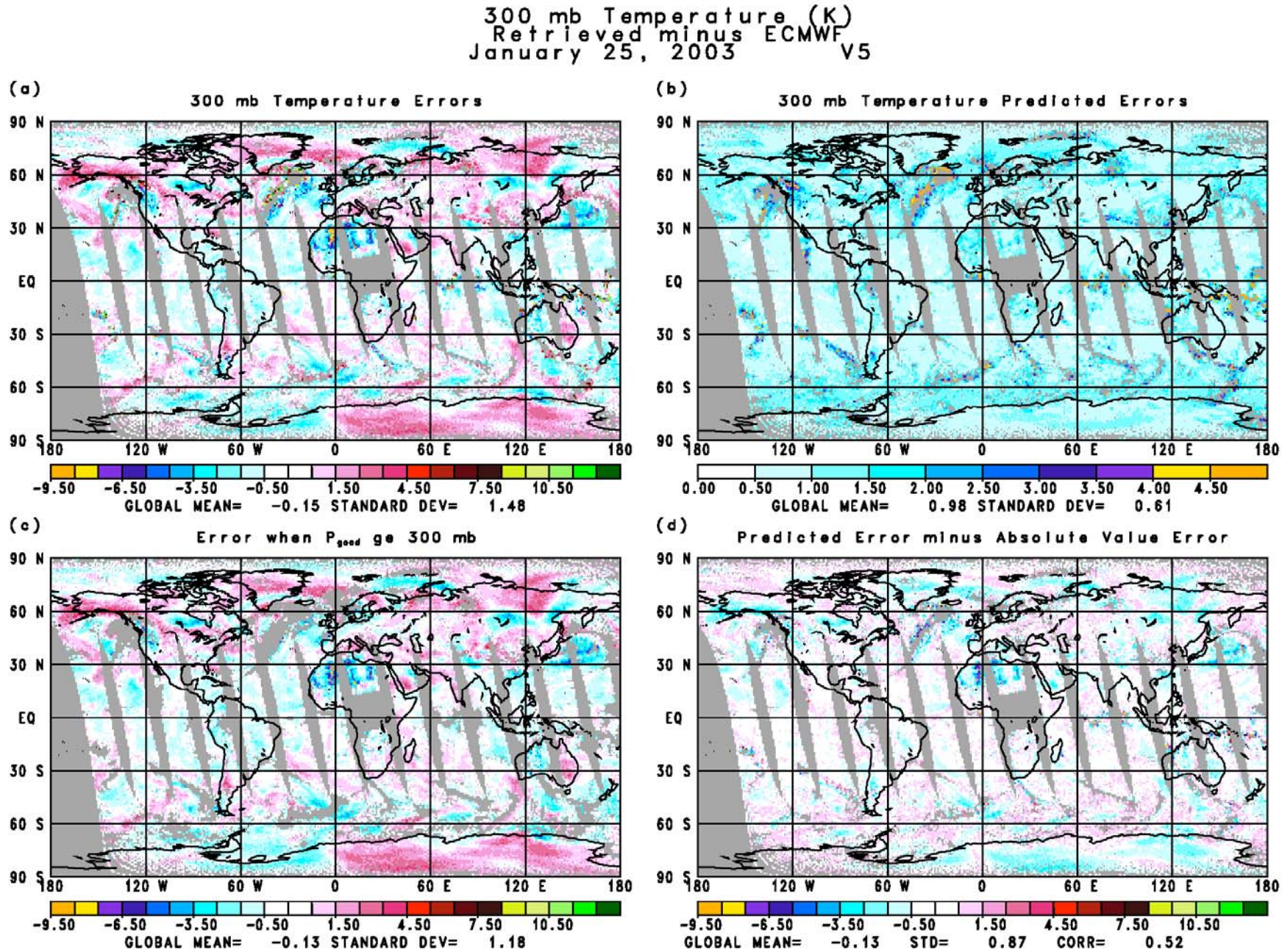
We have done forecast impact experiments with other thresholds: Medium and Tight

Assess trade-off between coverage and accuracy in data assimilation

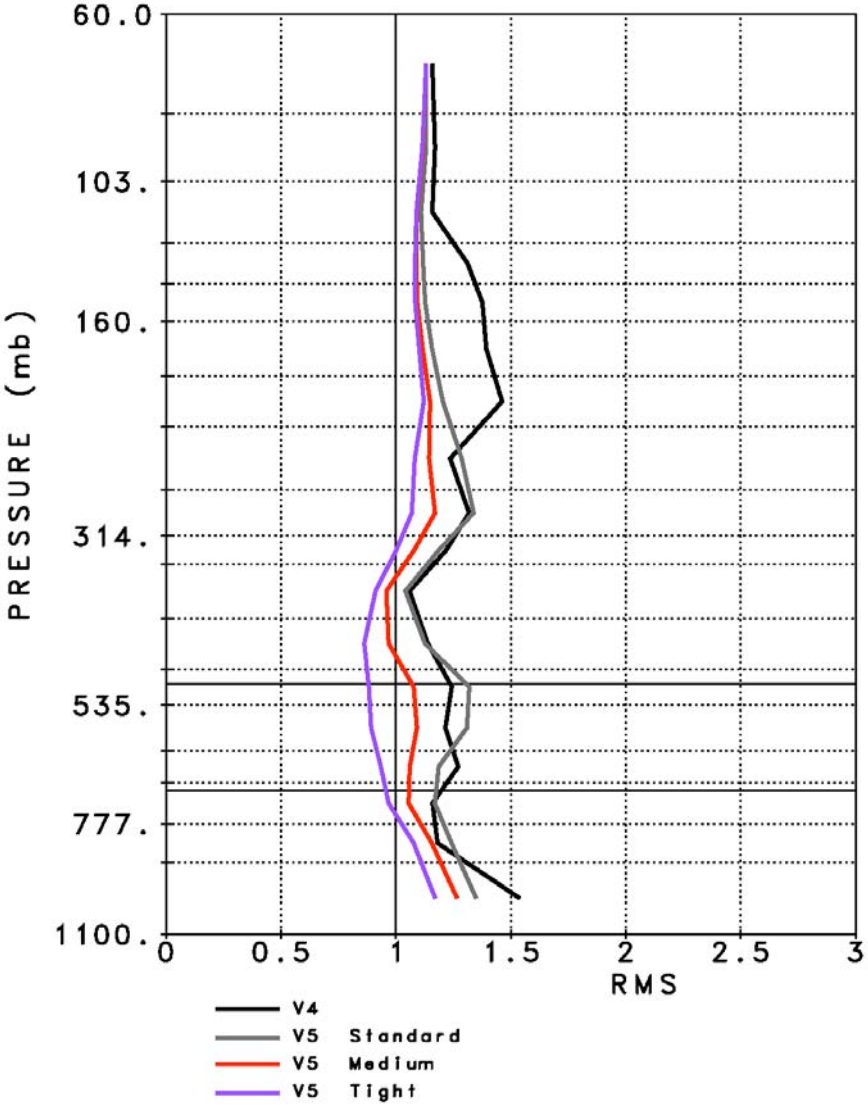
Temperature Profile Thresholds (K)

	Ocean			Land/Ice		
	ΔT_{70}	ΔT_{mid}	ΔT_{surf}	ΔT_{70}	ΔT_{mid}	ΔT_{surf}
Standard	1.75	1.25	2.25	2.25	2.0	2.0
Medium	1.75	1.0	1.75	1.75	1.0	2.0
Tight	1.75	0.75	1.75	1.75	0.75	1.75

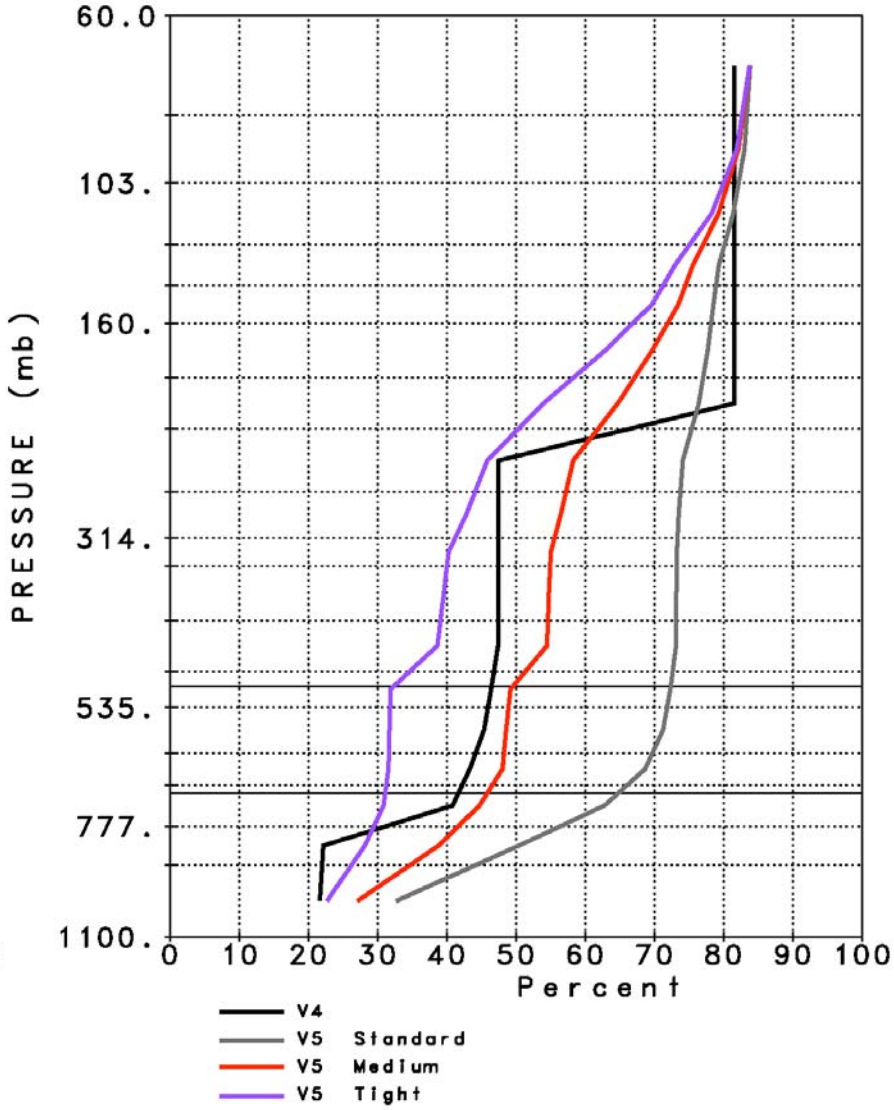

AIRS Cloud Top Pressure(mb)
Daytime 1:30 PM
January 25, 2003
90 N
60 N
30 N
EQ
30 S
60 S
90 S
180
120 W
60 W
0
60 E
120 E
180
lt 180 mb
180-310
310-440
440-560
560-680
680-800
gt 800 mb
0 40 80 % 40 80 40 80 40 80 40 80 40 80 40 80
GLOBAL MEAN= 45.18 STANDARD DEV= 31.22

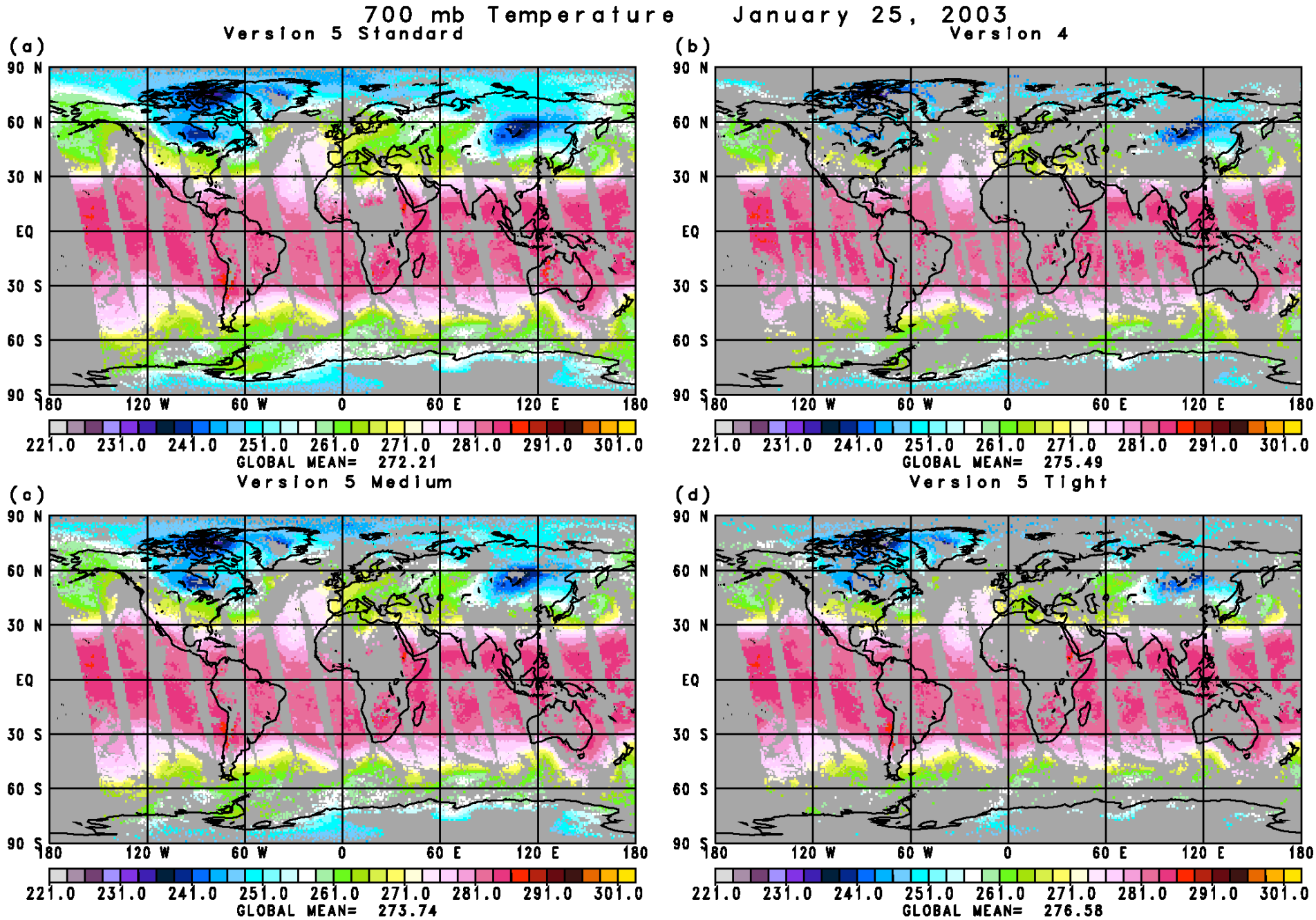

300 mb Temperature (K)
Retrieved minus ECMWF
January 25, 2003 V5
(a)
300 mb Temperature Errors
90 N
60 N
30 N
EQ
30 S
60 S
90 S
180
120 W
60 W
0
60 E
120 E
180
-9.50 -6.50 -3.50 -0.50 1.50 4.50 7.50 10.50
GLOBAL MEAN= -0.15 STANDARD DEV= 1.48
(b)
300 mb Temperature Predicted Errors
0.00 0.50 1.00 1.50 2.00 2.50 3.00 3.50 4.00 4.50
GLOBAL MEAN= 0.98 STANDARD DEV= 0.61
(c)
Error when P_{good} ge 300 mb
-9.50 -6.50 -3.50 -0.50 1.50 4.50 7.50 10.50
GLOBAL MEAN= -0.13 STANDARD DEV= 1.18
(d)
Predicted Error minus Absolute Value Error
-9.50 -6.50 -3.50 -0.50 1.50 4.50 7.50 10.50
GLOBAL MEAN= -0.13 STD= 0.87 CORR= 0.52

Layer Mean RMS Temperature (°C)
Global Differences from ECMWF
January 25, 2003
Global



Percent of all Cases Included
January 25, 2003
Global




700 mb Temperature January 25, 2003
(a)
Version 5 Standard
GLOBAL MEAN= 272.21
(b)
Version 4
GLOBAL MEAN= 275.49
(c)
Version 5 Medium
GLOBAL MEAN= 273.74
(d)
Version 5 Tight
GLOBAL MEAN= 276.58
90 N
60 N
30 N
EQ
30 S
60 S
90 S
180
120 W
60 W
0
60 E
120 E
180
221.0
231.0
241.0
251.0
261.0
271.0
281.0
291.0
301.0

Clear Column Radiances Error Estimates $\delta\hat{R}_i$

$$\hat{R}_i = \bar{R}_i + \sum_{j=1}^{9} \eta_j \left(R_{i,j} - \bar{R}_i \right)$$

If all η_j were perfect

$$\delta\hat{R}_i^{per} = \left[\left(\sum_{j=1}^{9} \frac{1}{9} \cdot \left(1 + \sum_{j=1}^{9} \eta_j' \right) - \eta_j \right)^2 \right]^{1/2} NE\Delta N_i = A \; NE\Delta N_i \approx \sum_{j=1}^{9} \left[\eta_j^2 \right]^{1/2} NE\Delta N_i$$

A is the channel noise amplification factor

Larger η's (more cloud clearing) results in more channel noise in $\hat{R}_i$

If channel i does not see clouds, we set all $\eta_j = 0$ $A = 1/3$ (noise reduction)

Errors in η_j will result in additional errors in $\hat{R}_i$, which are correlated from channel to channel

We set $\delta\hat{R}_i = \left(A \; NE\Delta N_i \right) + \left[\left(\sum_{k=1}^{6} \hat{M}_{i,k} \, \delta T(p_k) \right) \right] + \hat{M}_{i,7} \left[\delta W_{tot} / W_{tot} \right]$

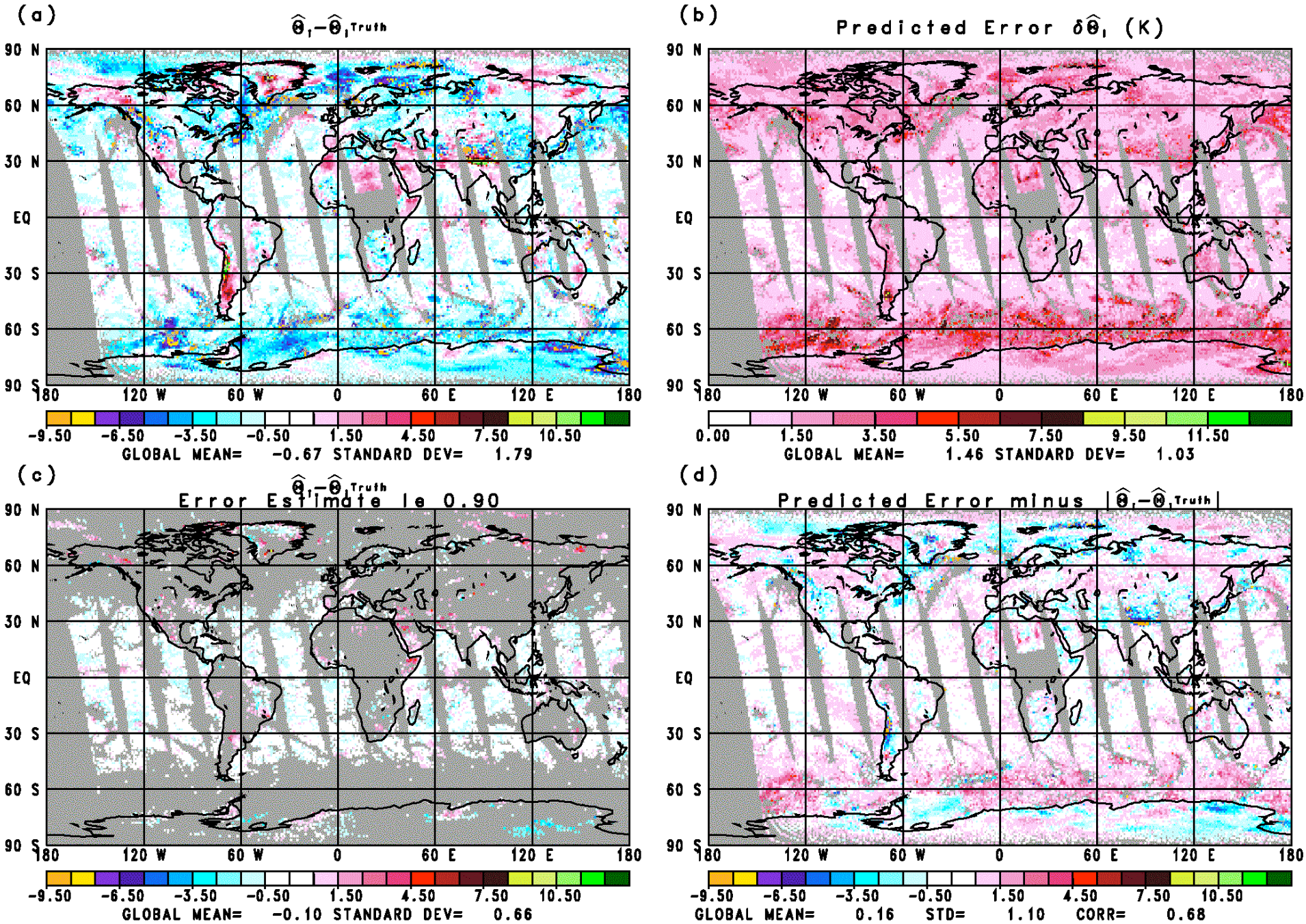
Where $\delta T(p_k)$ is the error estimate for $T(p_k)$ and $\delta W_{tot} / W_{tot}$ is the fractional error estimate for W_{tot}

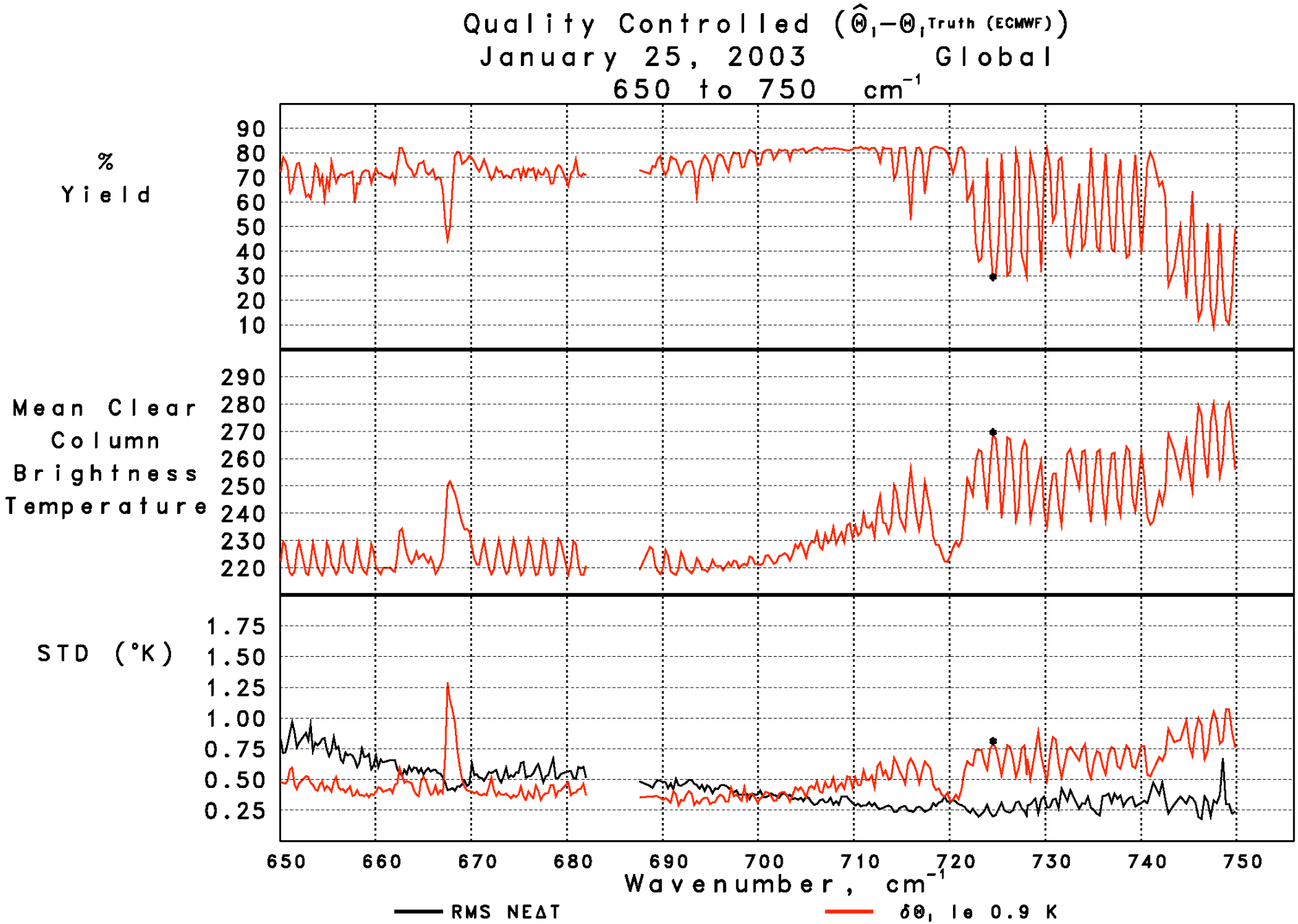
Coefficients of $\hat{M}$ are generated analogously to coefficients of M

Use $R_{i,CLR}$ computed from ECMWF as $\hat{R}_i^{truth}$

Error estimates are designed for use in temperature sounding regions only

Clear Column Brightness Temperature (K) minus ECMWF Truth
724.52 cm^{-1} Channel
January 25, 2003




Quality Controlled ($\hat{\Theta}_i - \Theta_{i\,Truth\ (ECMWF)}$)
January 25, 2003 Global
650 to 750 cm^{-1}
% Yield
90
80
70
60
50
40
30
20
10
Mean Clear Column Brightness Temperature
290
280
270
260
250
240
230
220
STD (°K)
1.75
1.50
1.25
1.00
0.75
0.50
0.25
650
660
670
680
690
700
710
720
730
740
750
Wavenumber, cm^{-1}
RMS NEΔT
$\delta\Theta_i$ le 0.9 K

First Set of Forecast Impact Tests

Experiments run with GSFC GEOS-5 Version 2 data assimilation system

- Same Version used for MERRA
- Forecasts run at 0.5° x 0.5° resolution
- Analysis using NCEP GSI analysis at 0.5° x 0.5° resolution

Data period covers January 1, 2003 - January 31, 2003

Control uses all data NCEP used operationally at that time

- Assimilates all satellite data but AIRS, including Aqua AMSU radiances

Control + AIRS adds V5.0 global quality controlled $T(p)$ retrievals

- Assimilated as if radiosonde data
- $\delta T(p)$ is used as the measurement error

27 independent forecasts run from each analysis

Forecasts verified against NCEP analysis

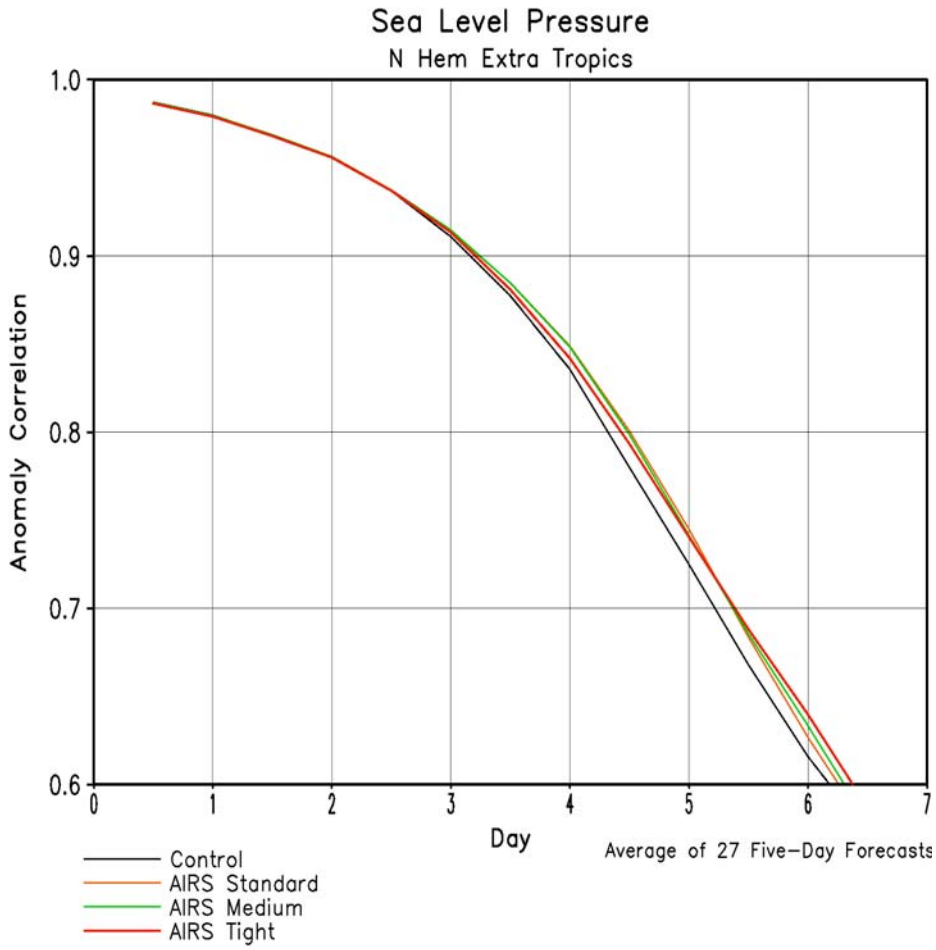
Experiment 1: Assessment of Trade-Off of Spatial Coverage and Overall Accuracy

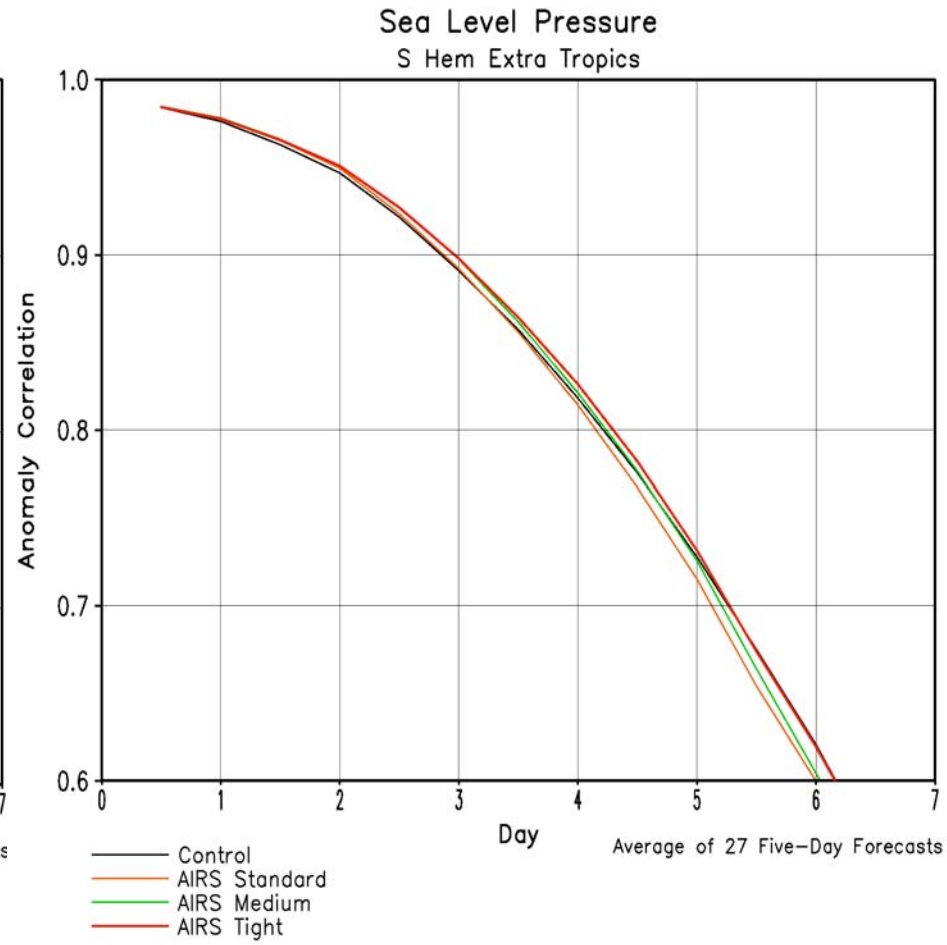
We compared forecasts from four assimilations

1a	Control
1b	AIRS V5 T(p) Standard QC
1c	AIRS V5 T(p) Medium QC
1d	AIRS V5 T(p) Tight QC

AIRS temperatures are assimilated down to p_g

Data assimilated in all three AIRS experiments is otherwise identical, except for p_g


Sea Level Pressure
N Hem Extra Tropics
Anomaly Correlation
1.0
0.9
0.8
0.7
0.6
0
1
2
3
4
5
6
7
Day
Average of 27 Five-Day Forecasts
Control
AIRS Standard
AIRS Medium
AIRS Tight


Sea Level Pressure
S Hem Extra Tropics
Anomaly Correlation
1.0
0.9
0.8
0.7
0.6
0
1
2
3
4
5
6
7
Day
Average of 27 Five-Day Forecasts
Control
AIRS Standard
AIRS Medium
AIRS Tight

Findings of Experiment 1 Assessed by SLP Anomaly Correlation

Northern Hemisphere Extra-tropics

All three AIRS data assimilation experiments improved SLP forecast skill significantly compared to the control

Northern hemisphere extra-tropics improvement in 5 day SLP forecast skill

- 5 hours for Tight QC, Medium QC, and Standard QC
- 5 hour improvement holds at 7 days for Tight QC
  - Improvement decreases at 7 days with looser QC

Southern Hemisphere extra-tropics

- Tight QC maintains control SLP forecast skill at 7 days

Medium QC and Standard QC degrade SLP forecast skill at 7 days

Tight QC performed significantly better than Standard QC which was optimized for climate

- Better overall accuracy with poorer spatial coverage
- Mid-latitude/tropical ocean spatial coverage is still extensive

Experiment 2: Test of The Importance of Assimilation of Tropospheric Temperatures

Motivation

Tony McNally at ECMWF stated that most of the impact of AIRS radiances on ECMWF analysis comes from 15μm CO_2 stratospheric sounding channels

We compared forecasts from four assimilations

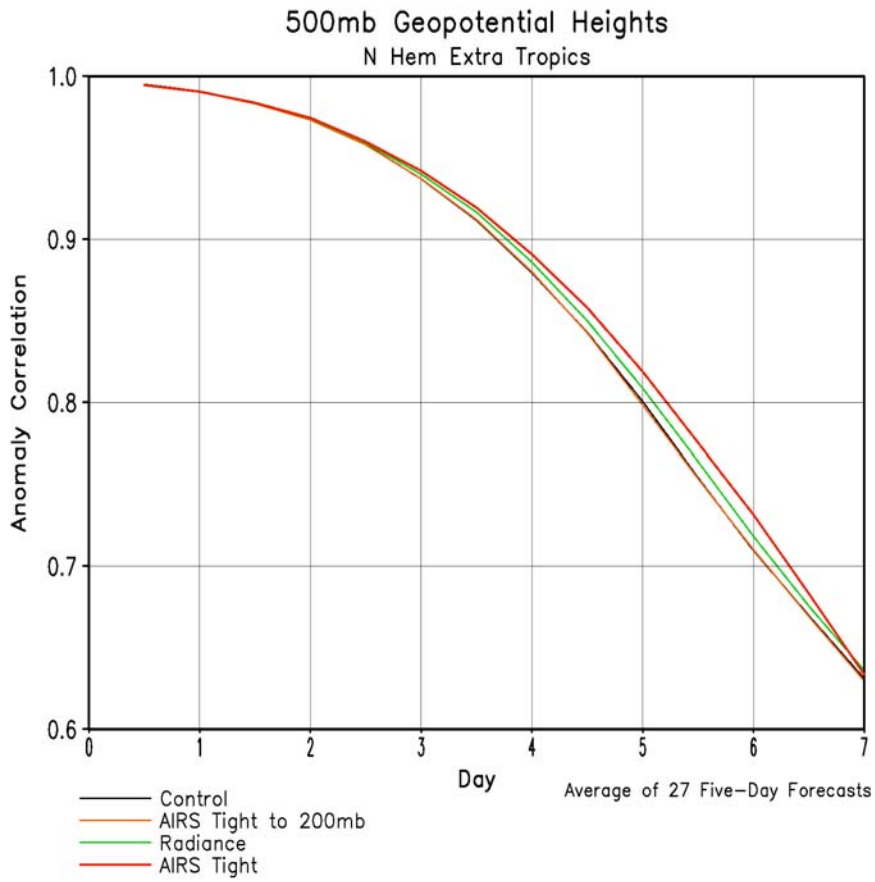
2a Control – same as 1a

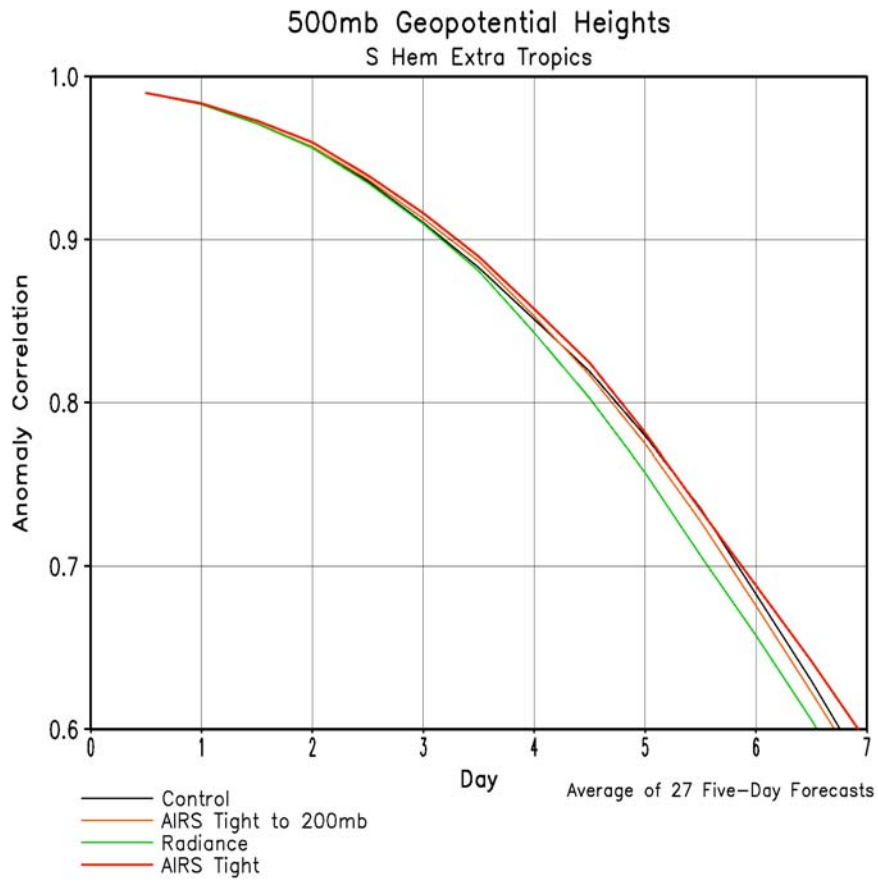
2b AIRS V5 T(p) Tight QC – same as 1d

2c AIRS V5 T(p) Tight QC but only down to 200 mb

2d AIRS radiance assimilation – uses primarily stratospheric AIRS radiance information

Assimilates only radiances unaffected by clouds


500mb Geopotential Heights
N Hem Extra Tropics
Anomaly Correlation
1.0
0.9
0.8
0.7
0.6
0
1
2
3
4
5
6
7
Day
Average of 27 Five-Day Forecasts
Control
AIRS Tight to 200mb
Radiance
AIRS Tight


500mb Geopotential Heights
S Hem Extra Tropics
Anomaly Correlation
1.0
0.9
0.8
0.7
0.6
0
1
2
3
4
5
6
7
Day
Average of 27 Five-Day Forecasts
Control
AIRS Tight to 200mb
Radiance
AIRS Tight

Findings of Experiment 2 - Assessed by 500 mb Height Anomaly Correlation

Northern Hemisphere Extra-tropics

Assimilation of AIRS temperature soundings with Tight QC gives 5 hour improvement in 5 day forecast skill

Assimilation of AIRS temperature soundings only down to 200 mb produced essentially no forecast impact

Most important T(p) information is coming from tropospheric temperatures determined under partial cloud cover

Assimilation of AIRS radiances had a small (2 hour) positive impact at 5 days

AIRS cloud free radiances contain some tropospheric information - but is sub-optimal

Southern Hemisphere Extra-tropics

Assimilation of AIRS temperature soundings with Tight QC improves 7 day forecast skill by 4 hours

Assimilation of AIRS temperatures down to 200 mb produces small negative impact

Assimilation of AIRS radiances produces a substantial negative impact

We are looking into the cause of this

New Forecast Impact Study in a More Recent Time Period

We recently conducted a forecast impact test for the time period 4/15/08 - 5/18/08

- Different season - transitional season
- More recent time period
- Contains cyclone Nargis - flooded Myanmar
- North Indian Ocean tropical cyclones are difficult to analyze
  - Errors in the location of cyclone centers are often 100 km in conventional analyses

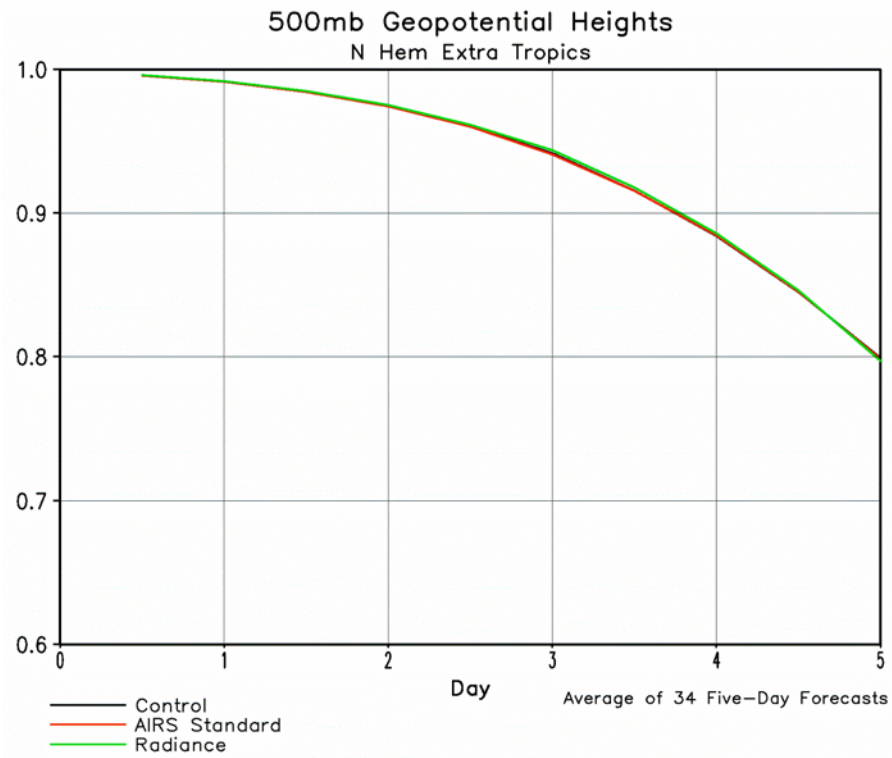
Experiments are analogous to those done previously

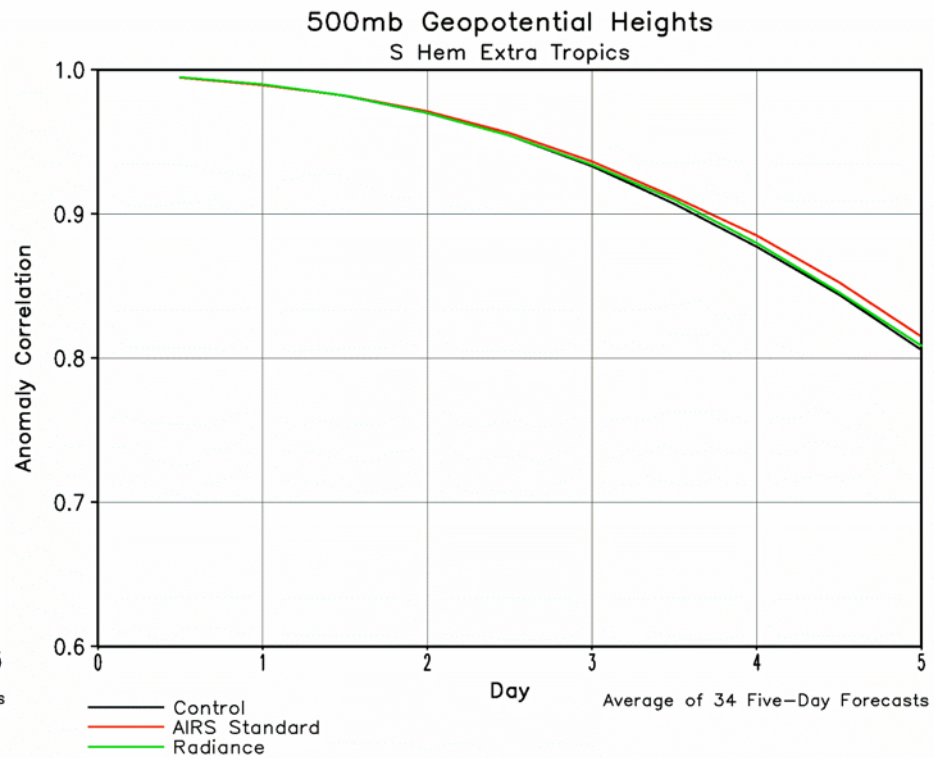
- Uses GEOS-5 Version 2 at 0.5° resolution
- AIRS T(p) obtained from Goddard DAAC - Standard QC only

Three data assimilation experiments run

- Control
- AIRS T(p) Standard QC
- AIRS radiances

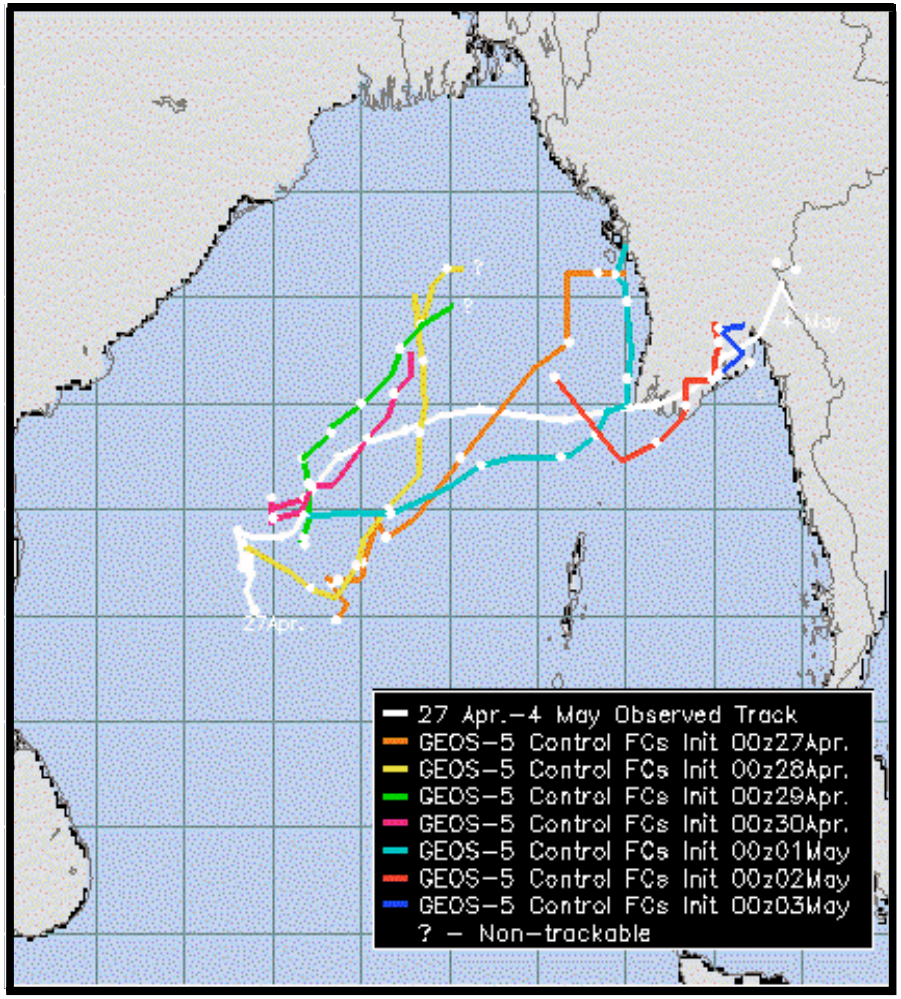
April 15 through May 18, 2008



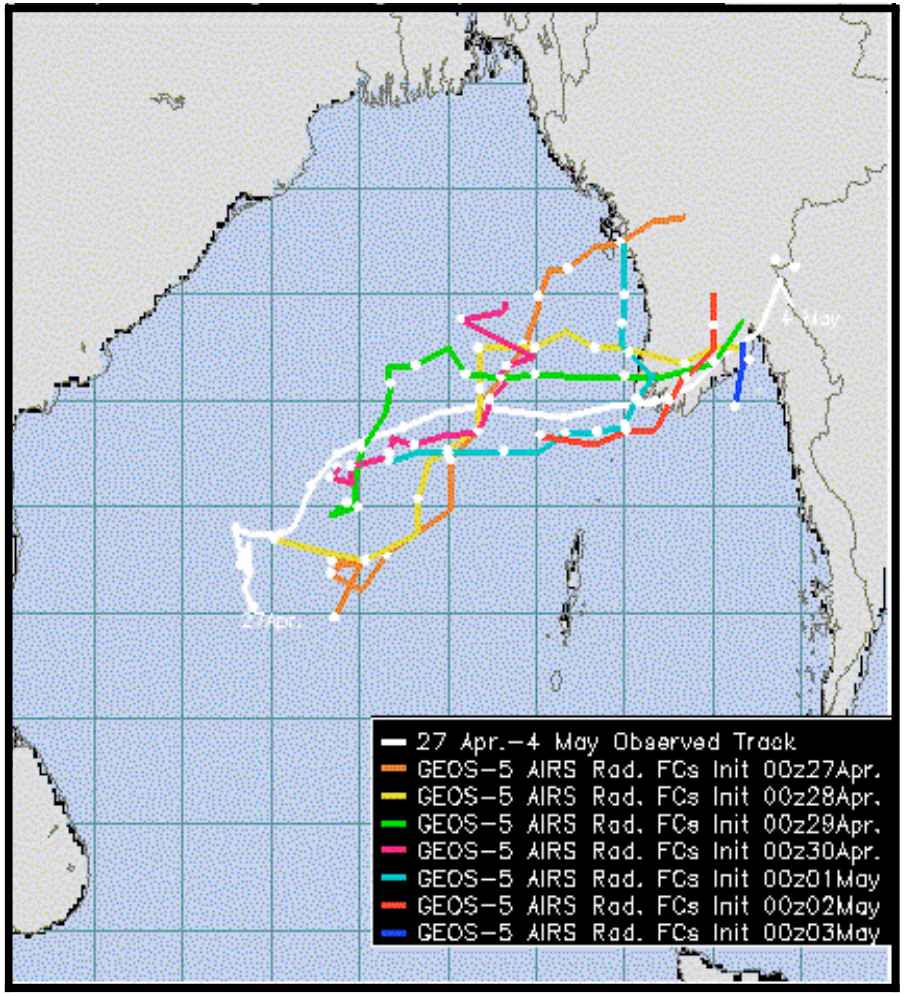


Bay of Bengal Nargis Cyclone Tracks

GEOS-5 Control Forecasts



AIRS Radiance Assimilation Forecasts



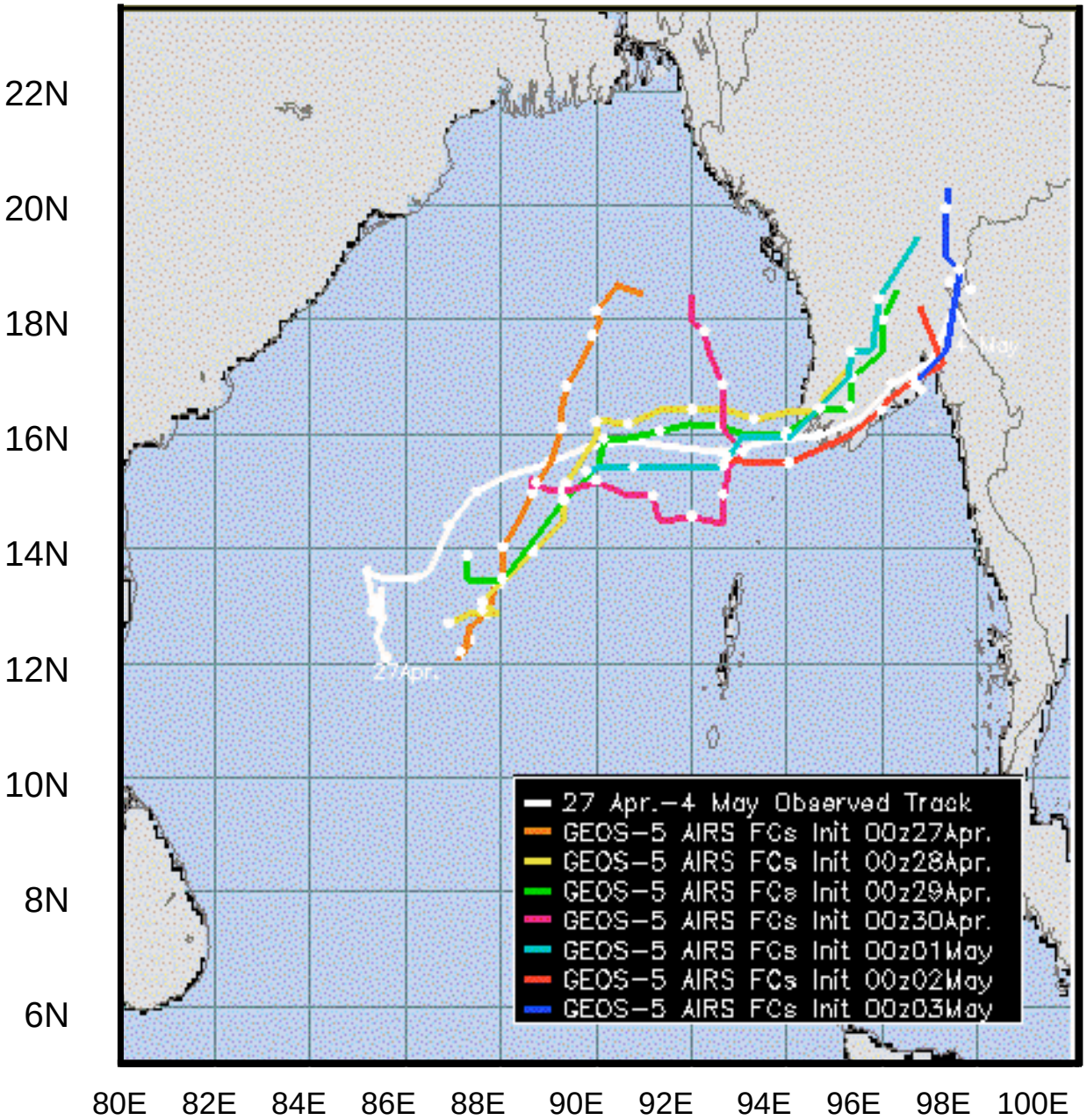
22N 20N 18N 16N 14N 12N 10N 8N 6N

80E 82E 84E 86E 88E 90E 92E 94E 96E 98E 100E

80E 82E 84E 86E 88E 90E 92E 94E 96E 98E 100E

Bay of Bengal Nargis Cyclone Tracks

AIRS T(p) Assimilation Forecasts



Findings of New Forecast Impact Test

Relative extra-tropical forecast impact results in Spring/Fall are comparable to Winter/Summer

- Assimilation of AIRS QC T(p) improves 5 day SH extra-tropics by 5 hours compared to control
- Assimilation of AIRS Radiances has much smaller forecast improvement

Assimilation of AIRS T(p) retrievals improves the analysis of weather systems in the tropics

- Analyses of Nargis storm location and confinement were improved significantly

Summary

AIRS Version 5 temperature profiles are determined primarily from 4.2 μm CO_2 band

Allows for accurate retrievals under most cloud conditions

Soundings are flagged as good down to a pressure p_g using QC methodology

Data assimilation experiments were done using NCEP GSI analysis at 0.5° x 0.5° resolution

Assimilation of Quality Controlled AIRS T(p) significantly improves forecast skill

Most of the improvement results from assimilation of tropospheric T(p) under partial cloud cover

Assimilation of observed AIRS radiances performed poorer than assimilation of T(p)

Further experiments are planned to optimize T(p) QC for data assimilation

Working with Lars-Peter Riishojgaard, we plan pseudo-operational test at NCEP assimilating AIRS T(p) in place of AIRS R_i as done operationally

We plan to test assimilation of $\hat{R}_i$ using error estimate QC at GSFC

Should perform better than assimilation of R_i when $\hat{R}_i$ QC is optimized

NASA