Using Principal Component Analysis (PCA) of IASI radiances to filter noise and generate retrievals.

Mitchell D. Goldberg, Zhaohui Cheng , Walter Wolf,
Lihang Zhou, and Chris Barnet

NOAA/NESDIS
Center for Satellite Applications and Research

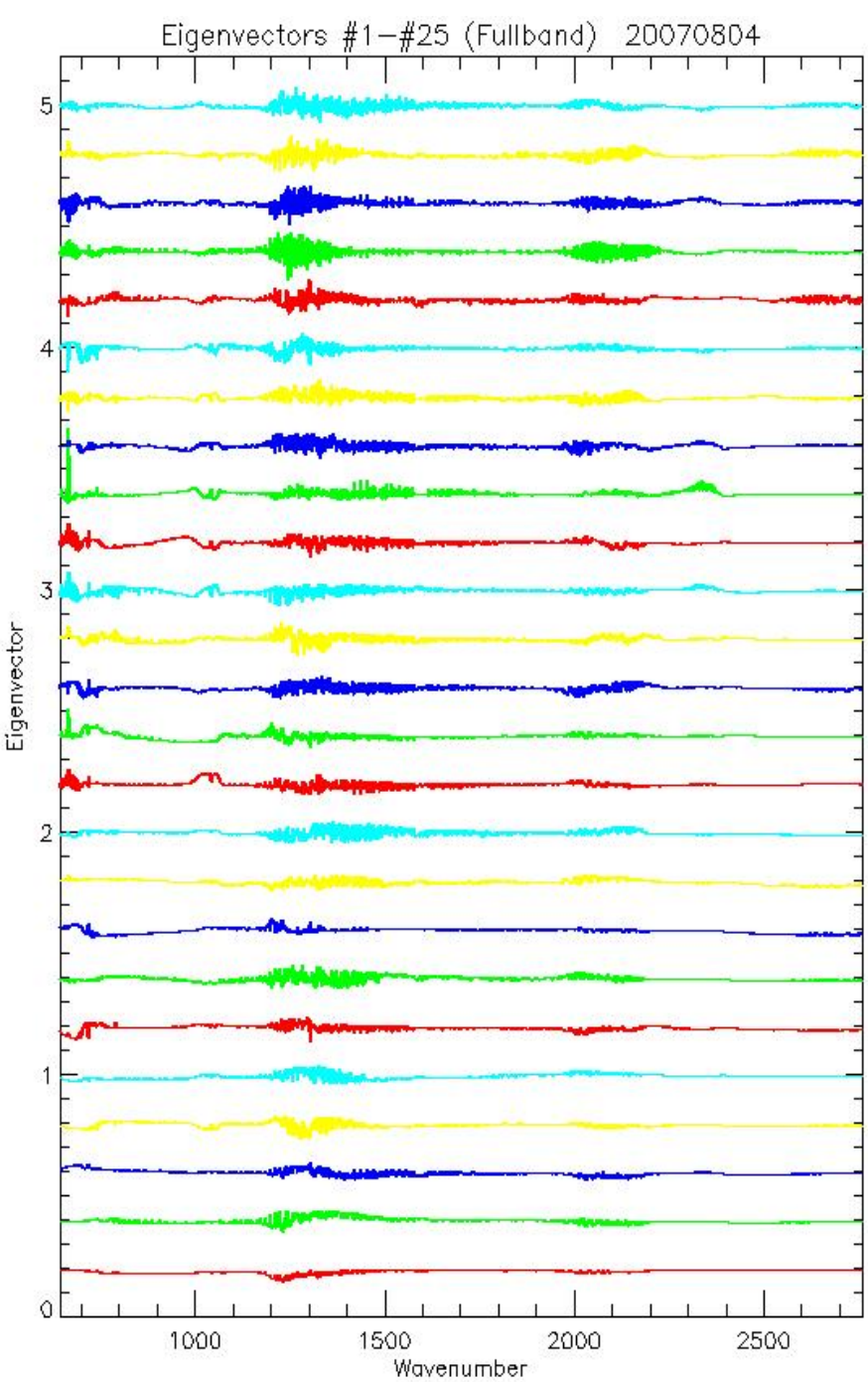
1st IASI Conference, Anglet, France

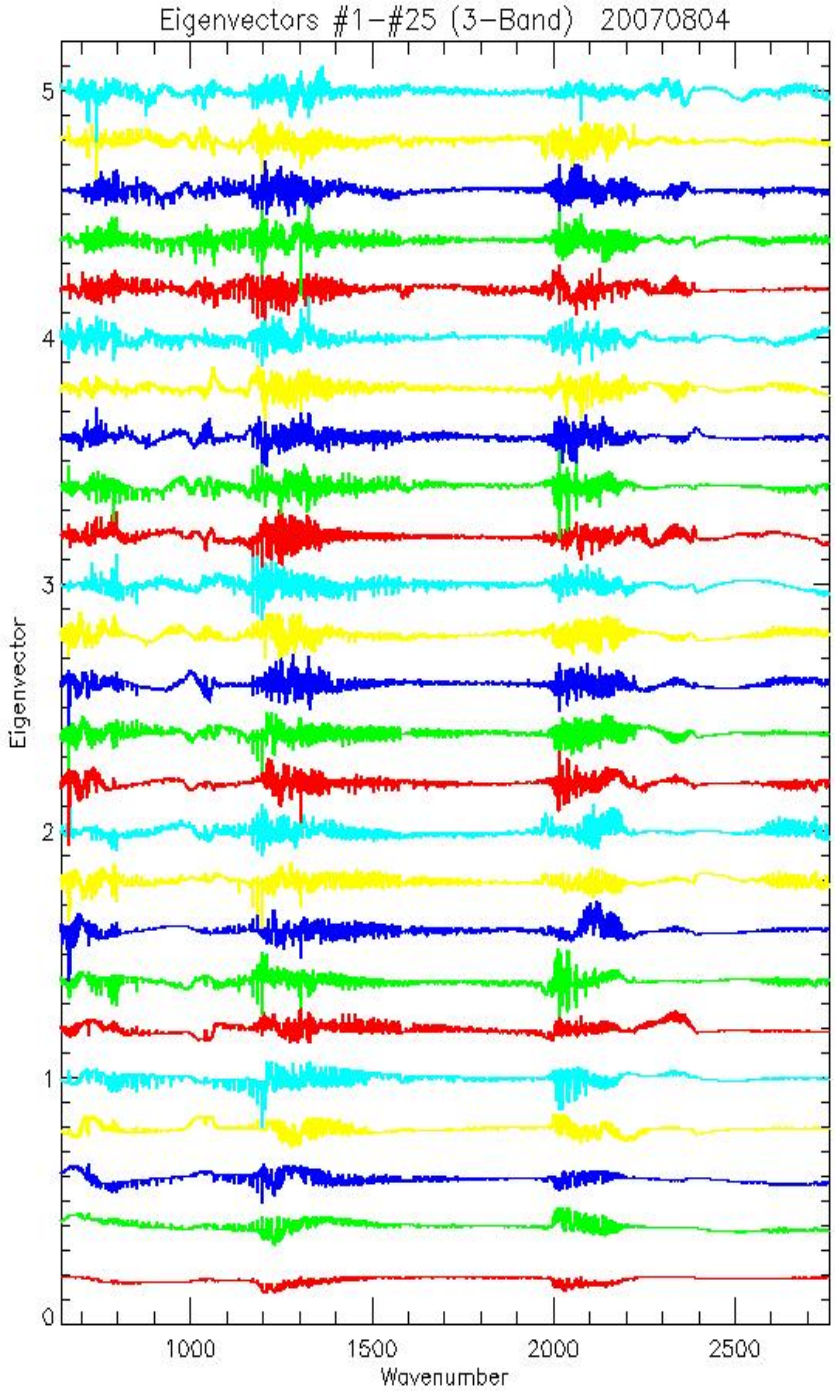
IASI PCA: Motivation

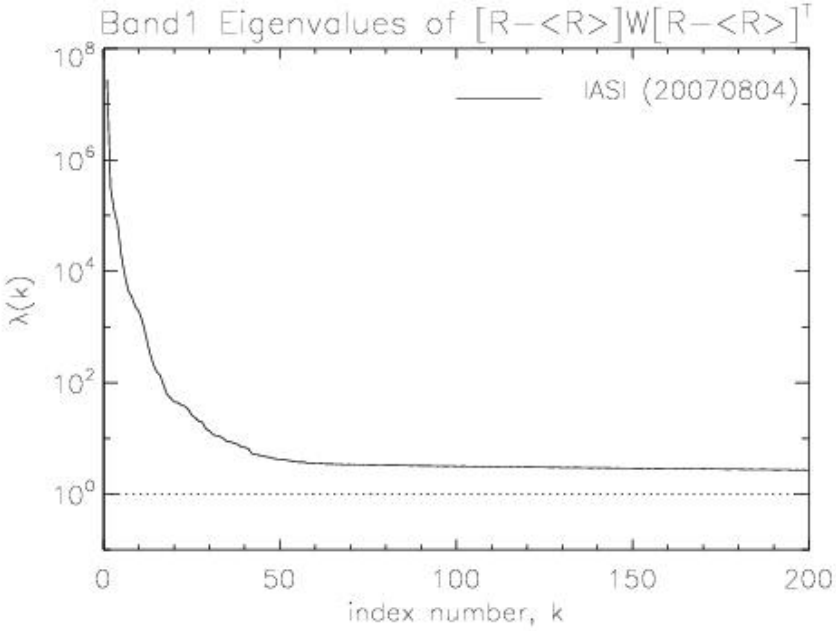
- IASI's 8461 channels can be well represented by relatively few empirical orthogonal functions (EOFs), also called eigenvectors or principle components.
- Each IASI spectra can be expressed as a linear function of these EOF's by a unique set of coefficients. These coefficients are also called principal component scores (PCS).
- We found that no more than 150 PCS are needed to reconstruct all the radiances within the noise, so there are potential important applications, (*e.g*. data compression, noise estimation, etc.).
- Reconstructed radiances can be used to estimate instrument noise in Earth scenes and detect anomalies by comparing reconstructed with the original spectra.
  - we can detect 2 sigma, 3 sigma events.
- We found there is no advantage of granule eigenvectors versus global eigenvectors, except for very extreme events such as volcanic eruptions which are currently not represented in the global eigenvector training set (*i.e*., training needs to be performed over larger ensemble of geophysical conditions).
- PCS used as regression predictors provide very fast retrievals. The accuracy of the regression is excellent in the free troposphere (away from the surface) and in the stratosphere. Improvements by a physical retrieval is generally realized in the lower troposphere, near the surface and for water vapor.

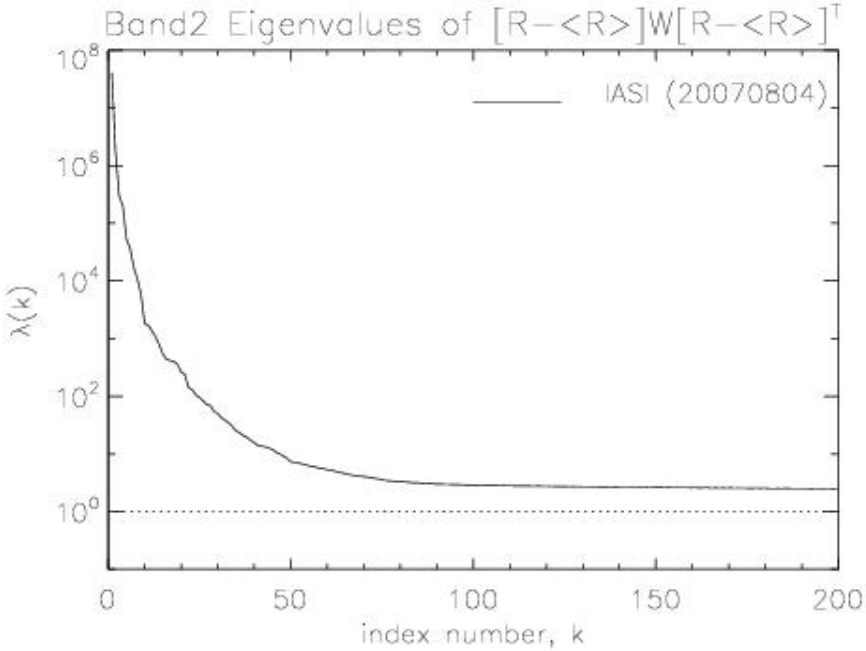
Generation of EOFs

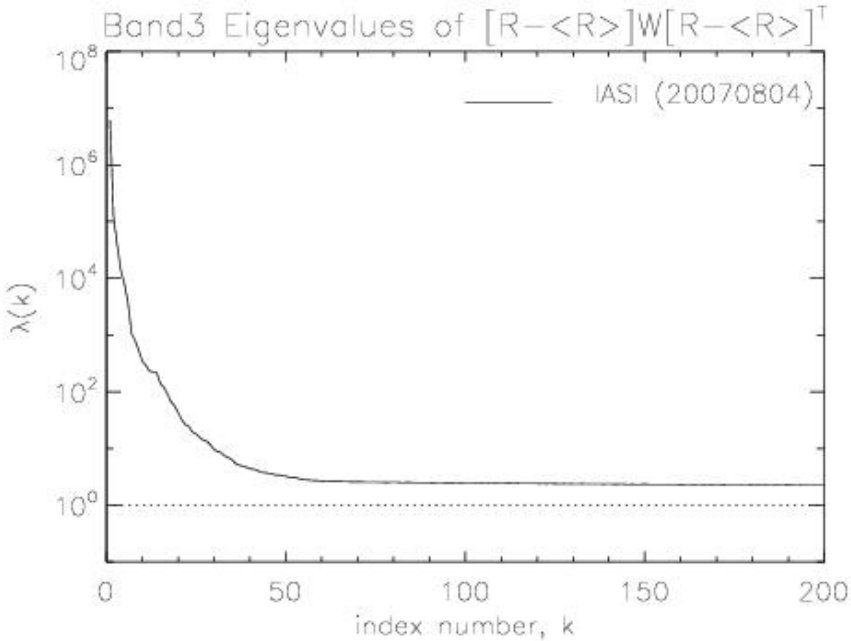
- 3 bands together (used in regression)
- Separate Bands (comparison of reconstruction scores from all band versus individual can be used to monitor long term stability)
  - 645- 1210 cm^{-1}
  - 1210 – 2000 cm^{-1}
  - 2000 – 2760 cm^{-1}
- We generate eigenvectors for all of the above
- Eigenvectors are generated from one single day of globally distributed IASI spectra – August 4, 2007.
- We plan to update after finding potential outliers (volcanic events, etc). We have not found any so far.

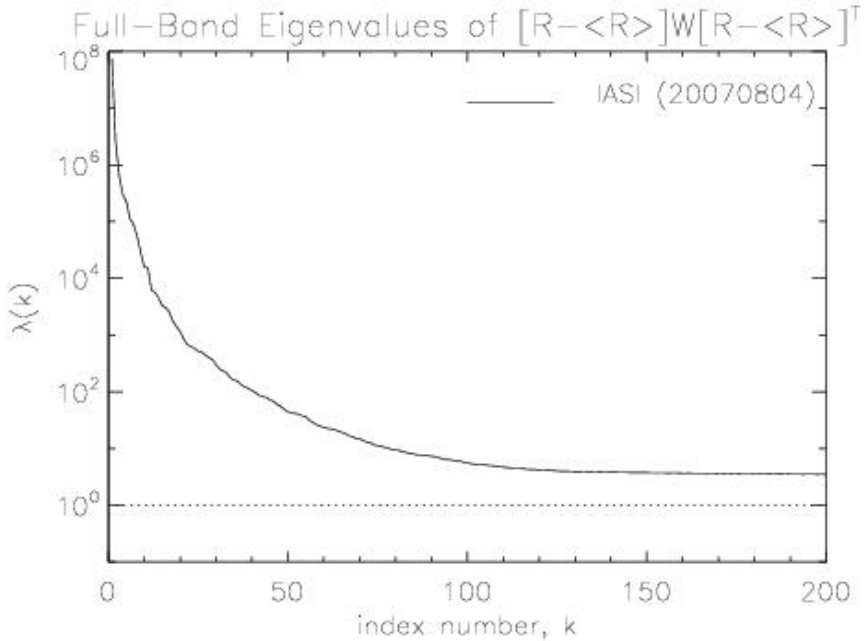

Eigenvectors #1-#25 (Fullband) 20070804
Eigenvector
Wavenumber
0
1
2
3
4
5
1000
1500
2000
2500


Eigenvectors #1-#25 (3-Band) 20070804
Eigenvector
Wavenumber
0
1
2
3
4
5
1000
1500
2000
2500


Band1 Eigenvalues of [R−<R>]W[R−<R>]T
IASI (20070804)
λ(k)
index number, k


Band2 Eigenvalues of [R−<R>]W[R−<R>]T
IASI (20070804)
λ(k)
index number, k

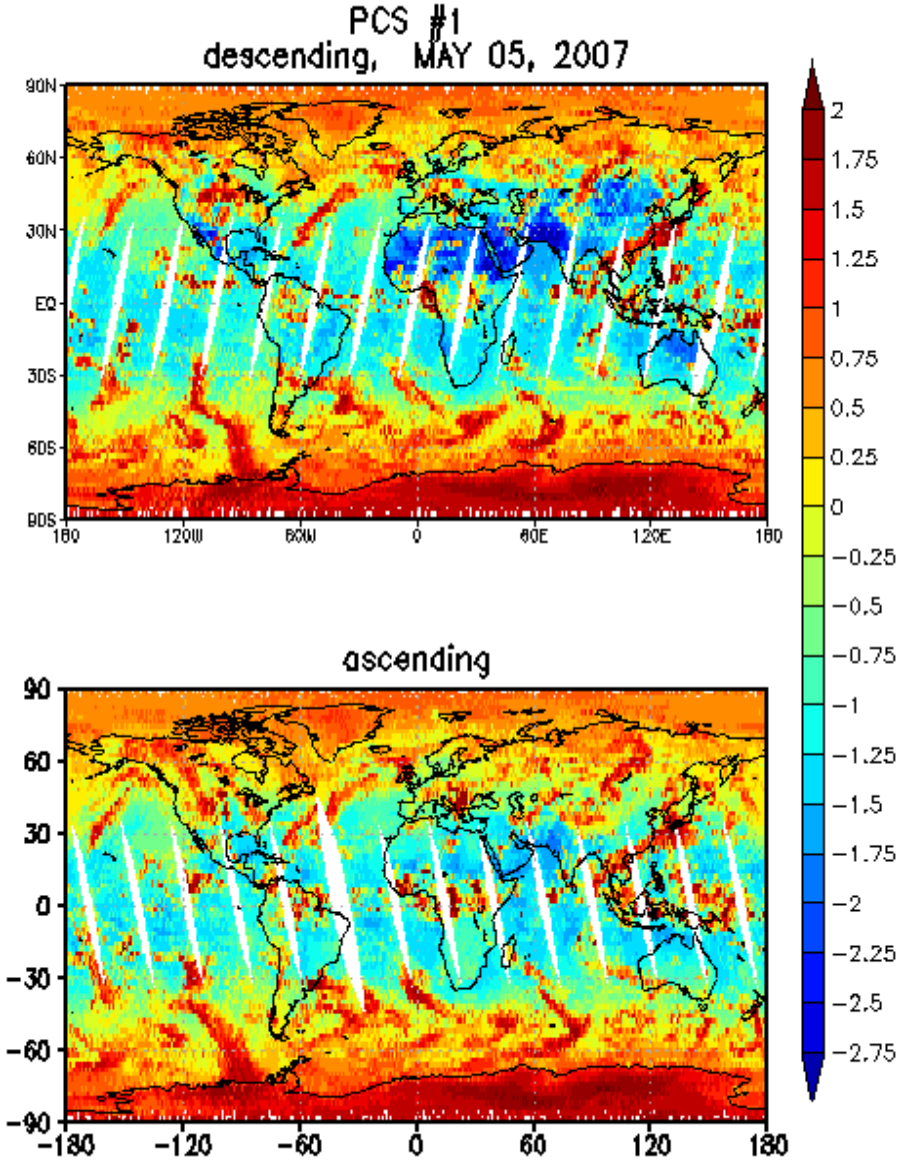

Band3 Eigenvalues of [R−<R>]W[R−<R>]T
IASI (20070804)
λ(k)
index number, k

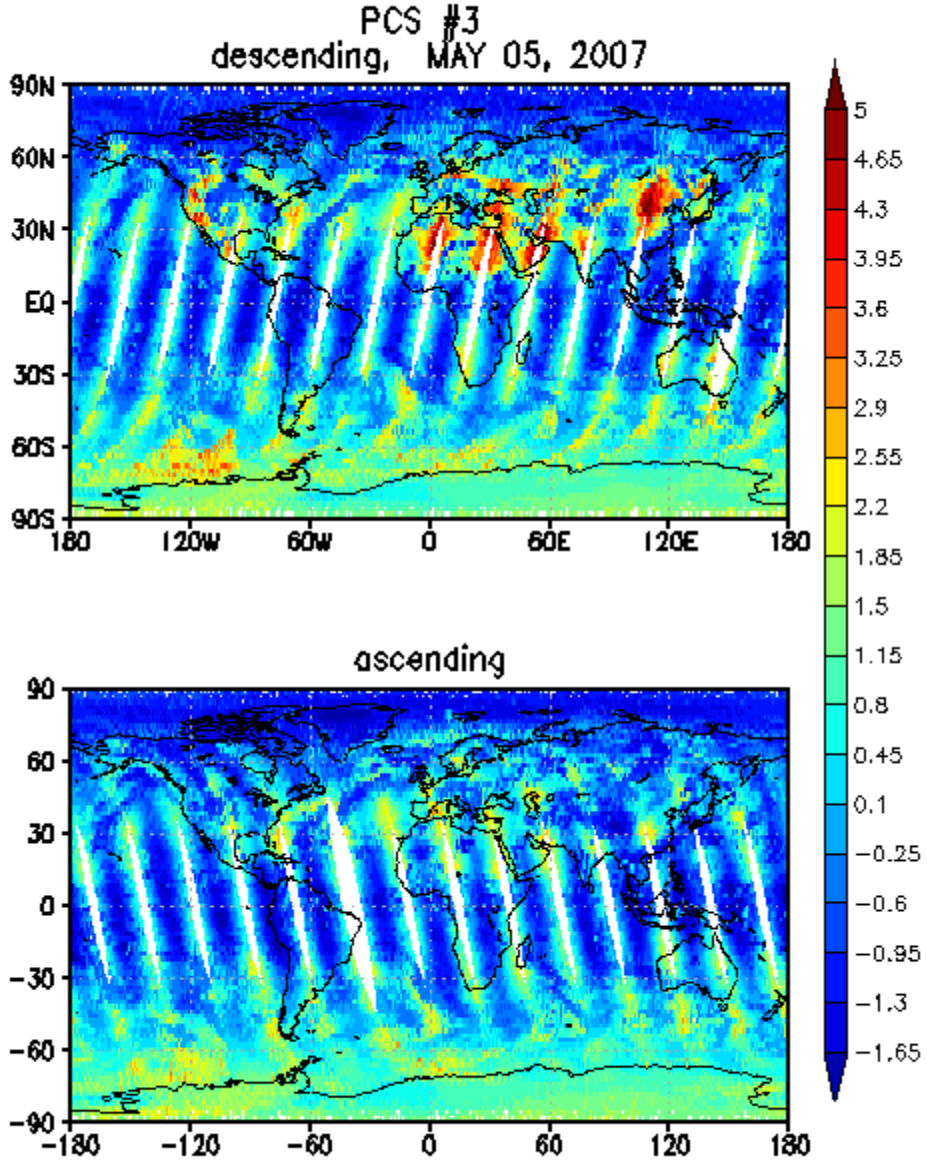

Full−Band Eigenvalues of [R−<R>]W[R−<R>]T
IASI (20070804)
λ(k)
index number, k

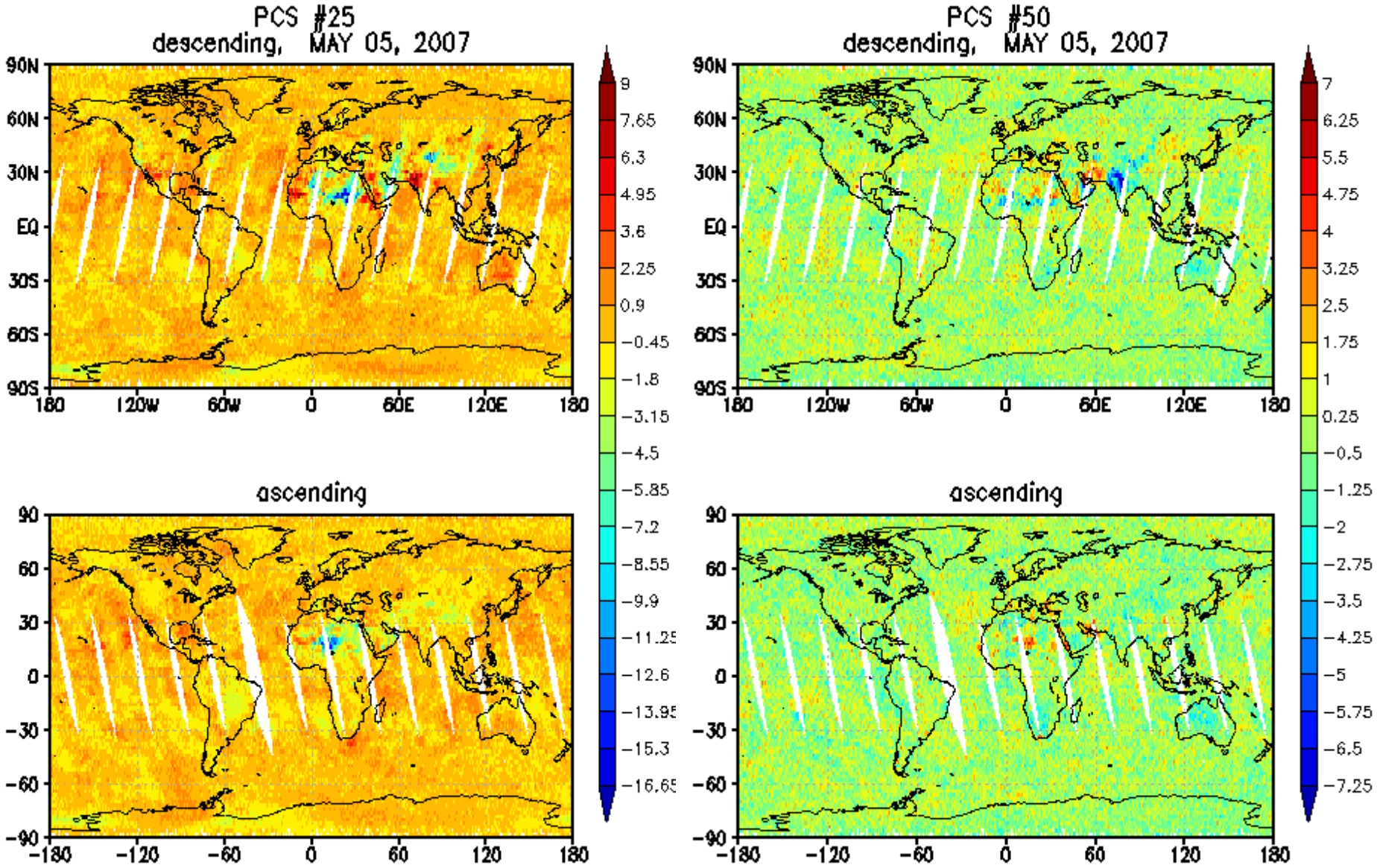
How to determine the number of EOF's to use?

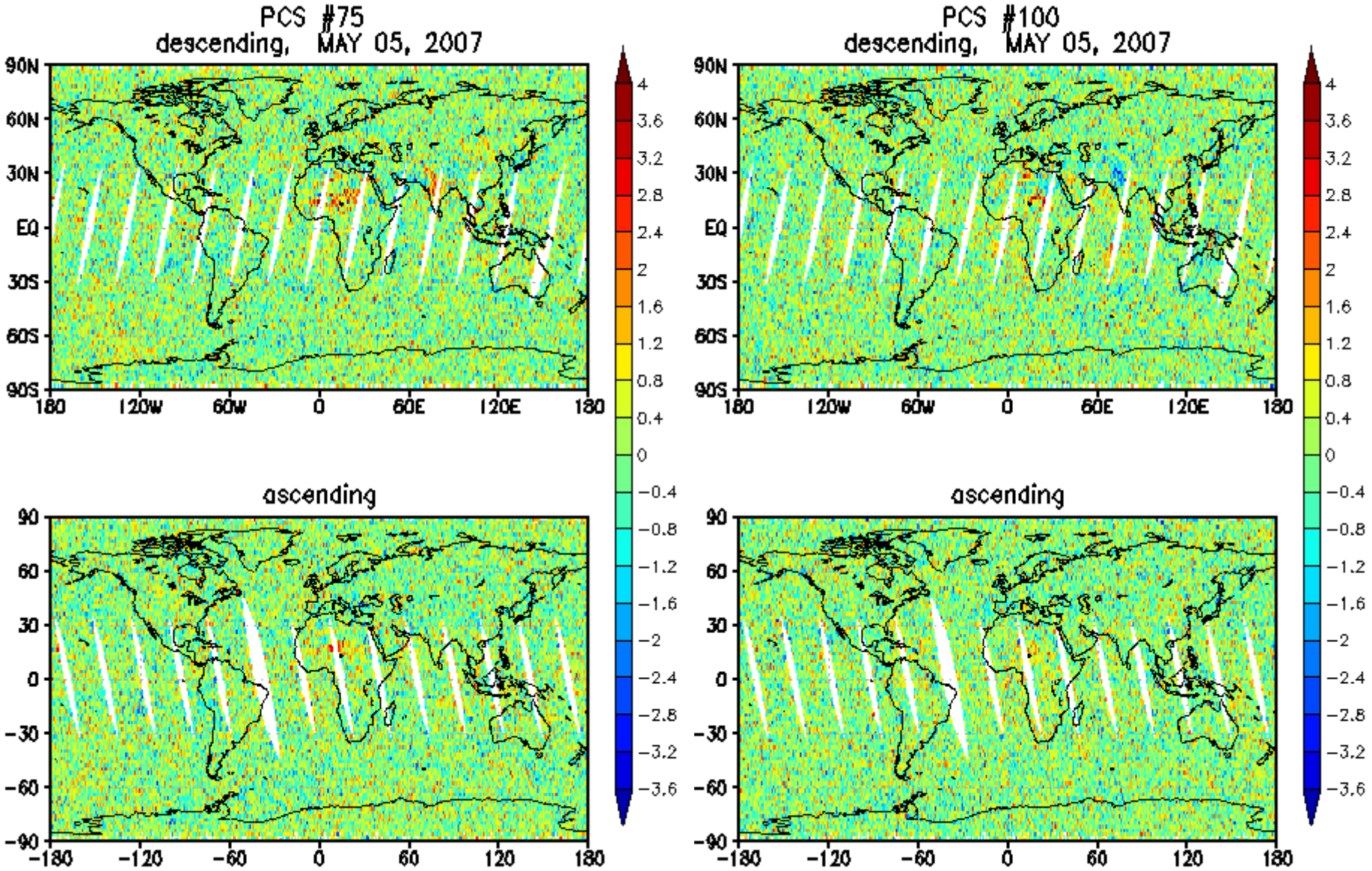
- Examine eigenvalues
- Examine global maps of PCS
- Verify that the spatial patterns in PCS are not correlated.
  - This always results in using more eigenvectors than just relying only on examination of the eigenvalues.

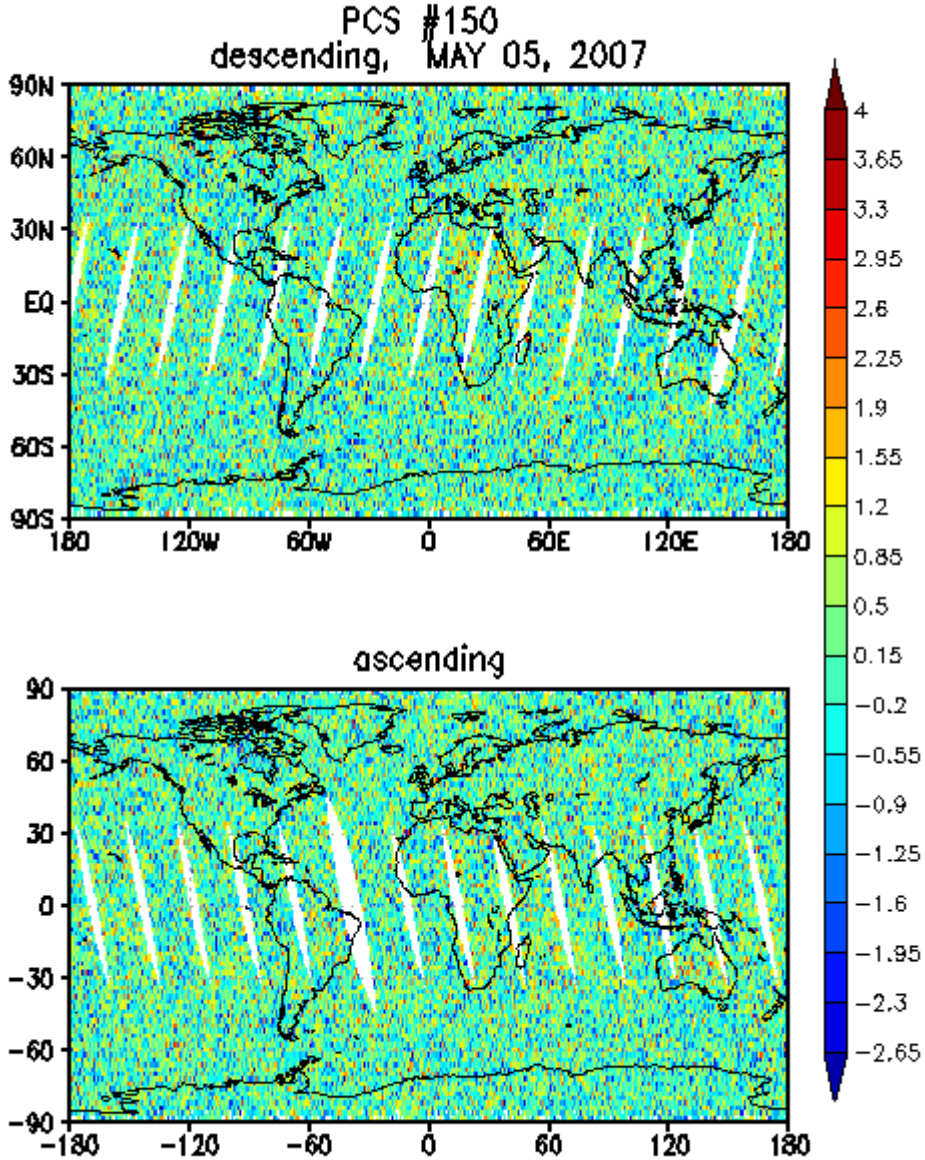
Examples of spatial patterns (geophysical variability) in significant" PCS's



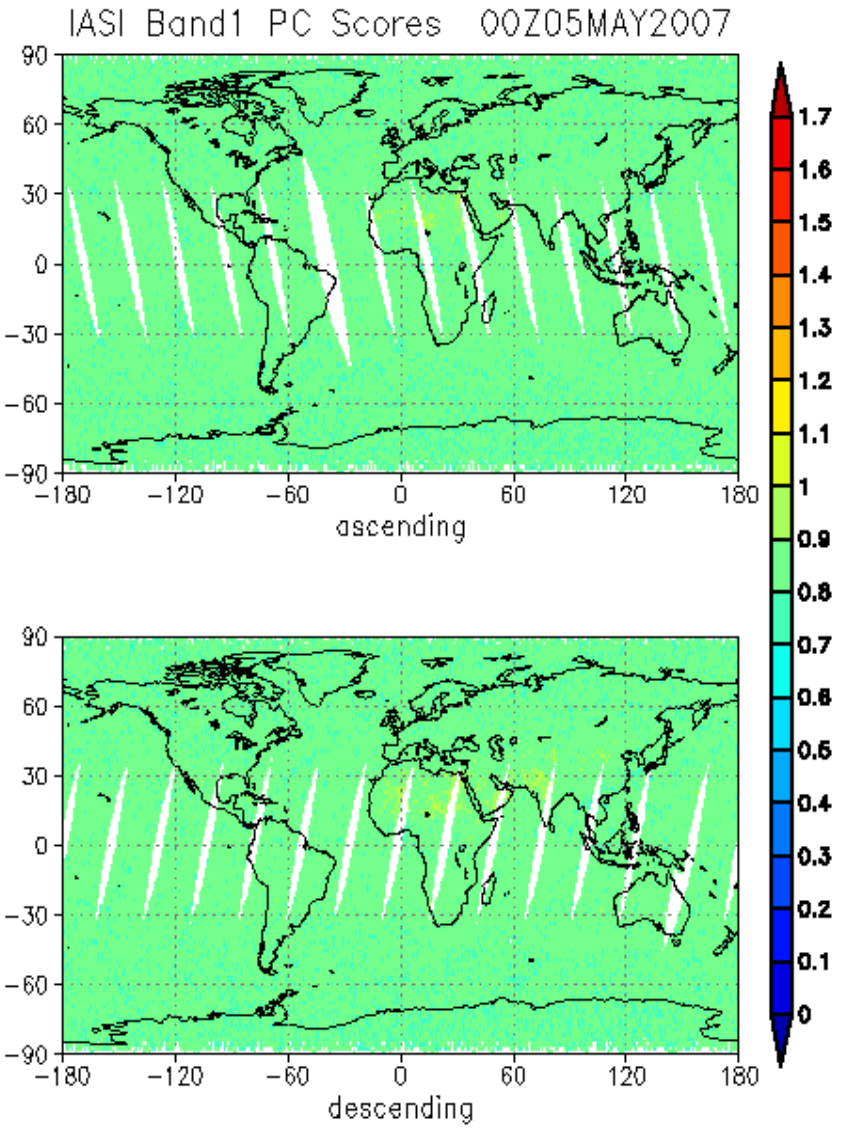


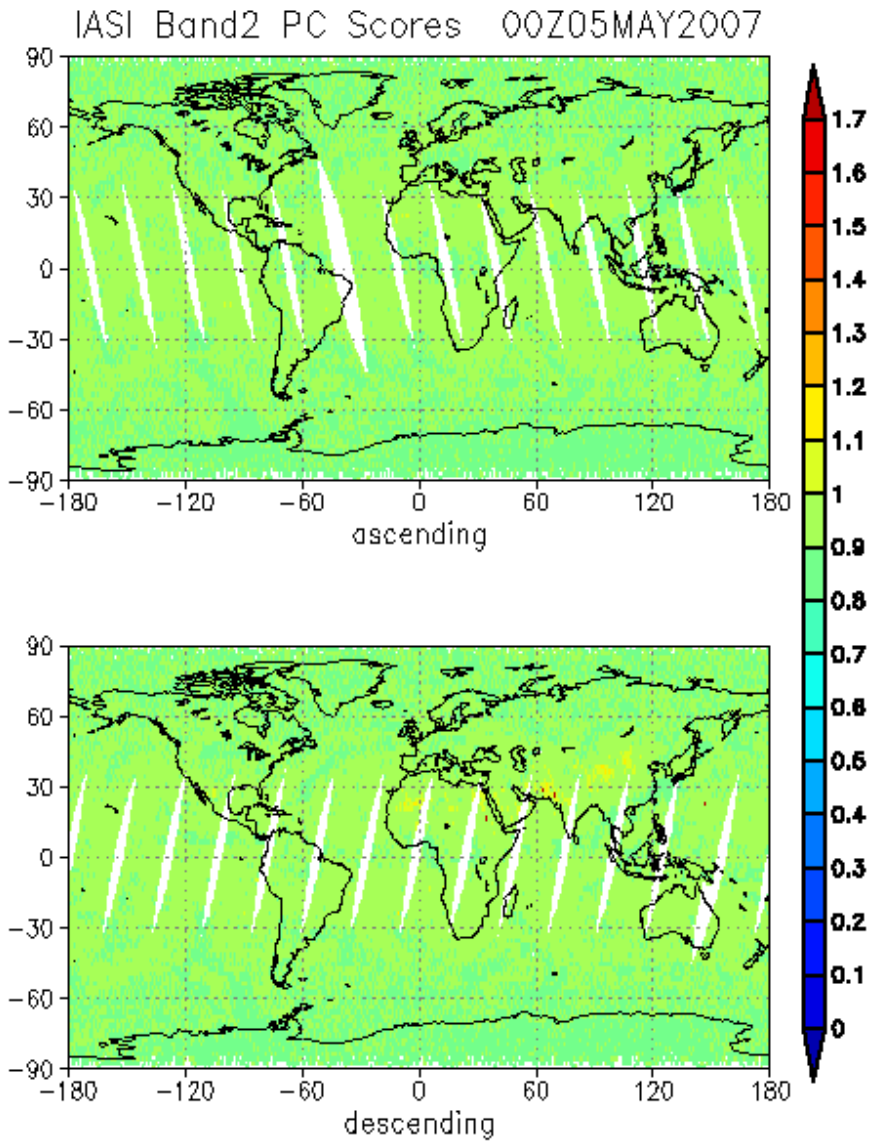

PCS #25
descending, MAY 05, 2007
90N
60N
30N
EQ
30S
60S
90S
180
120W
60W
0
60E
120E
180
ascending
90
60
30
0
-30
-60
-90
-180
-120
-60
0
60
120
180
9
7.65
6.3
4.95
3.6
2.25
0.9
-0.45
-1.8
-3.15
-4.5
-5.85
-7.2
-8.55
-9.9
-11.25
-12.6
-13.95
-15.3
-16.65
PCS #50
descending, MAY 05, 2007
ascending
7
6.25
5.5
4.75
4
3.25
2.5
1.75
1
0.25
-0.5
-1.25
-2
-2.75
-3.5
-4.25
-5
-5.75
-6.5
-7.25


PCS #75
descending, MAY 05, 2007
PCS #100
descending, MAY 05, 2007
ascending
ascending
90N
60N
30N
EQ
30S
60S
90S
180
120W
60W
0
60E
120E
180
90
60
30
0
-30
-60
-90
-180
-120
-60
0
60
120
180
4
3.6
3.2
2.8
2.4
2
1.6
1.2
0.8
0.4
0
-0.4
-0.8
-1.2
-1.6
-2
-2.4
-2.8
-3.2
-3.6

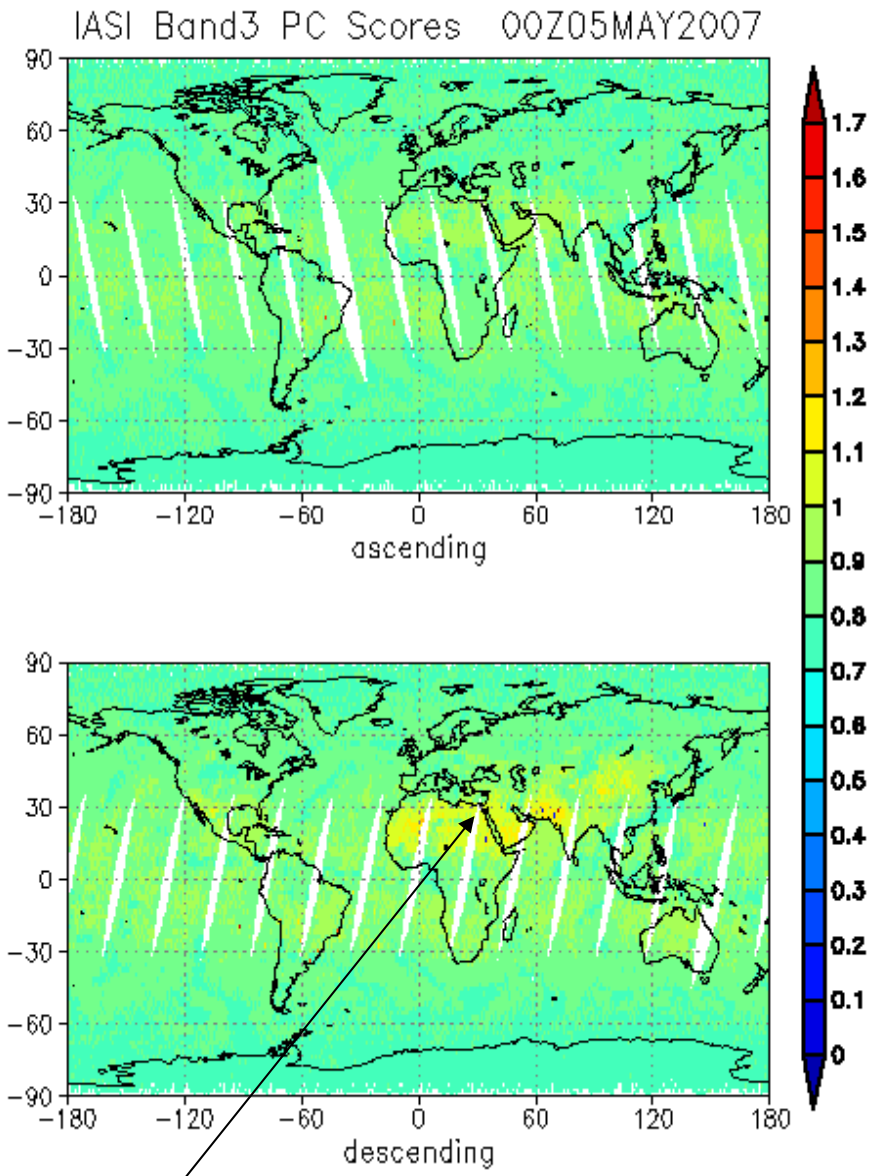

PCS #150
descending, MAY 05, 2007
90N
60N
30N
EQ
30S
60S
90S
180
120W
60W
0
60E
120E
180
ascending
90
60
30
0
-30
-60
-90
-180
-120
-60
0
60
120
180
4
3.65
3.3
2.95
2.6
2.25
1.9
1.55
1.2
0.85
0.5
0.15
-0.2
-0.55
-0.9
-1.25
-1.6
-1.95
-2.3
-2.65

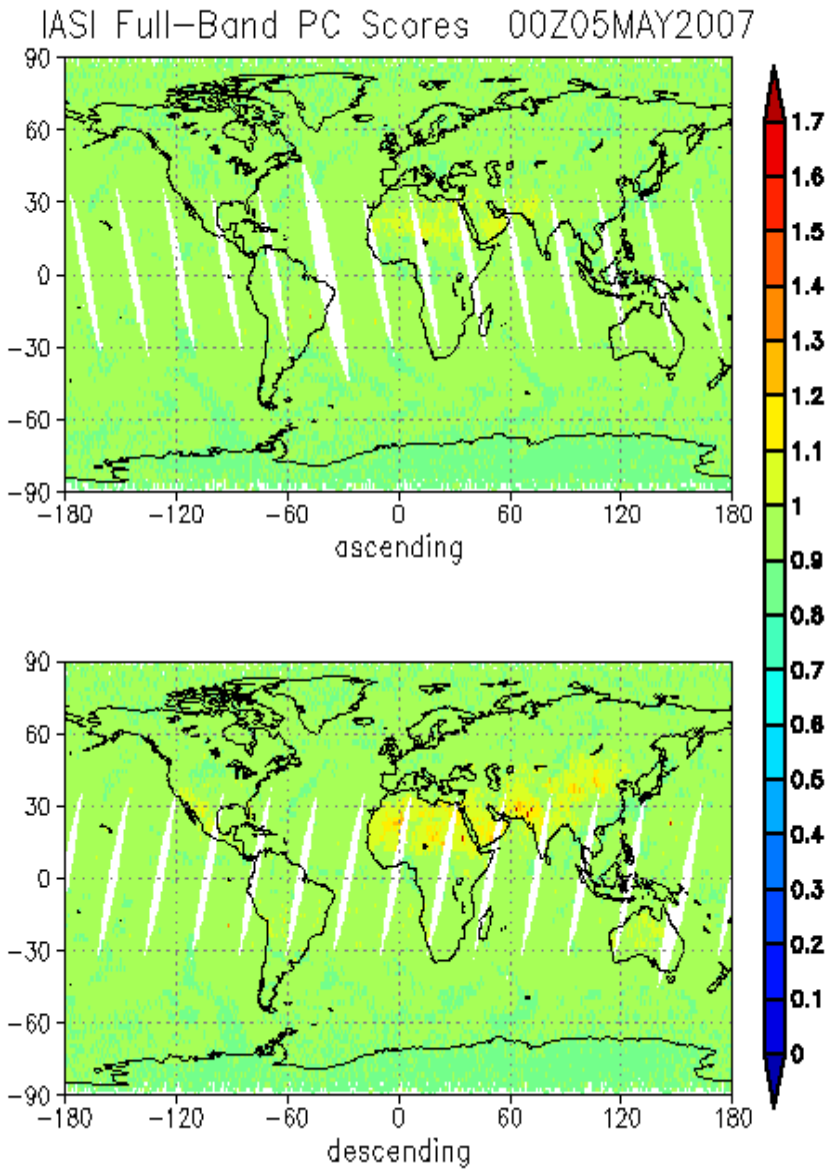
Reconstruction Scores (RMS of Obs-Reconstruction)
< 1 = radiances are reconstructed within noise level





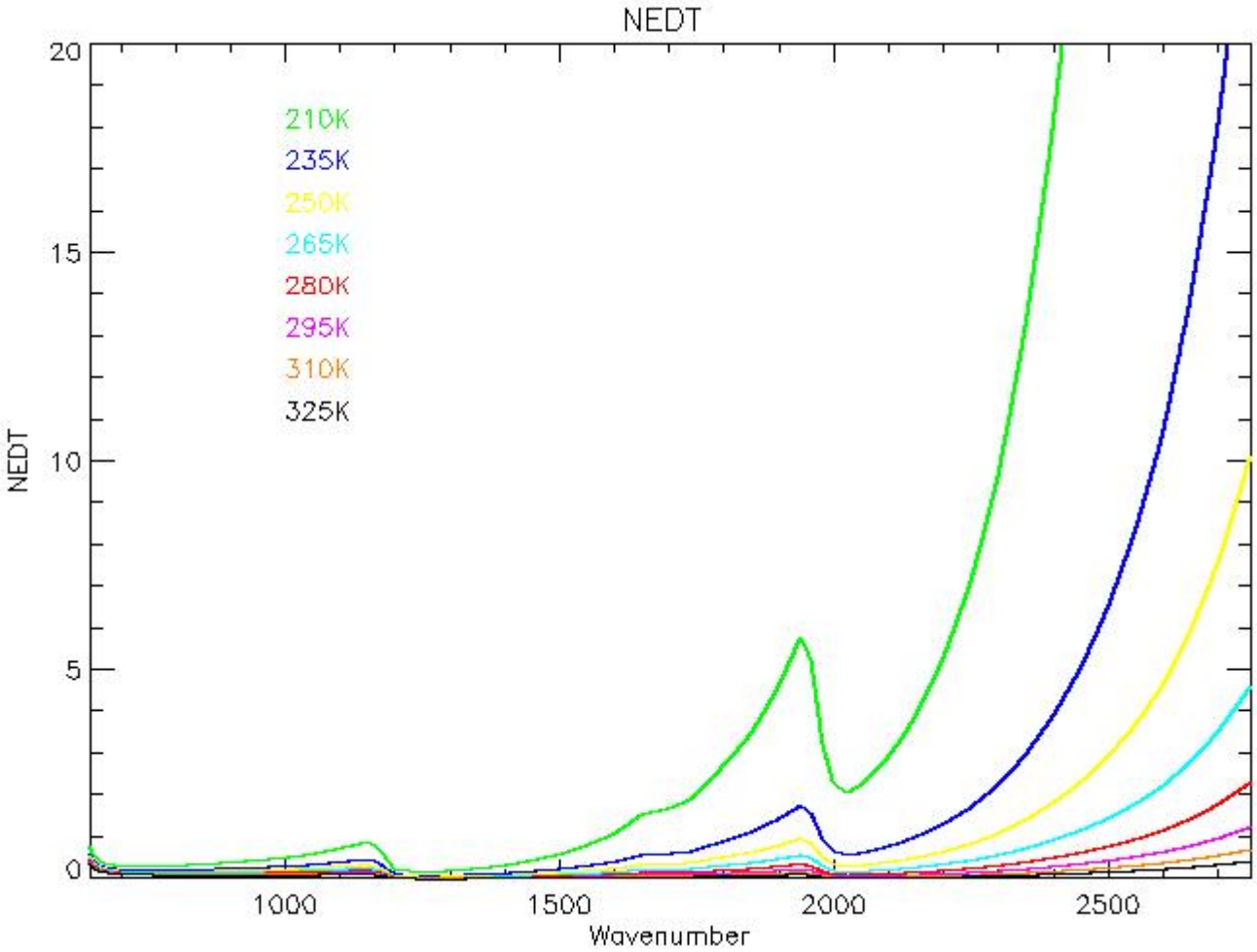
Reconstruction Scores (RMS of Obs-Reconstruction)
< 1 = radiances are reconstructed within noise level





10 -20% higher over hot ground, but NEDT is still very low for high temperatures

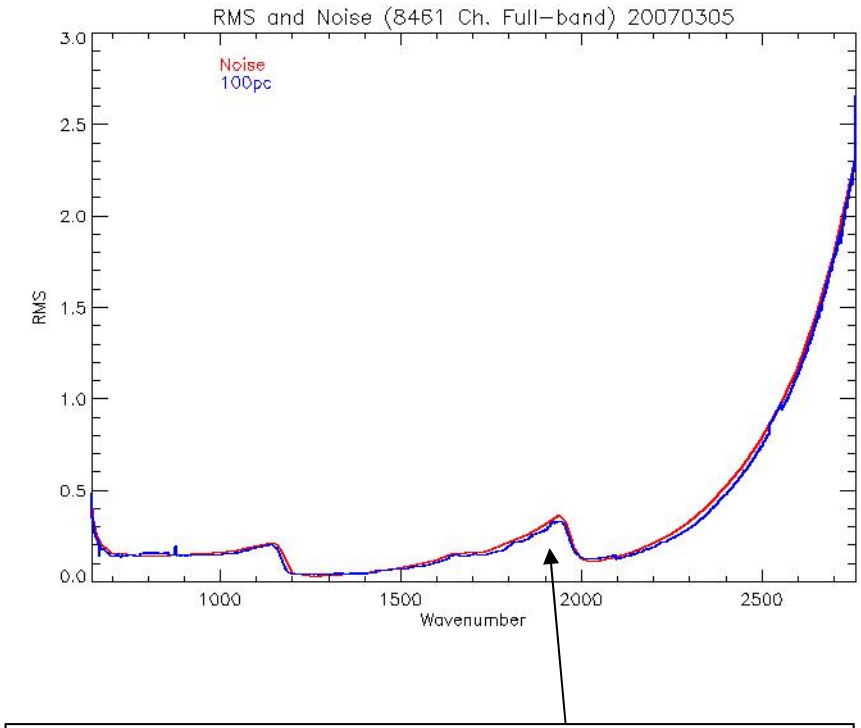
Reminder that noise is very nonlinear with temperature and wavenumber



Eigenvector Analysis for Noise Reduction

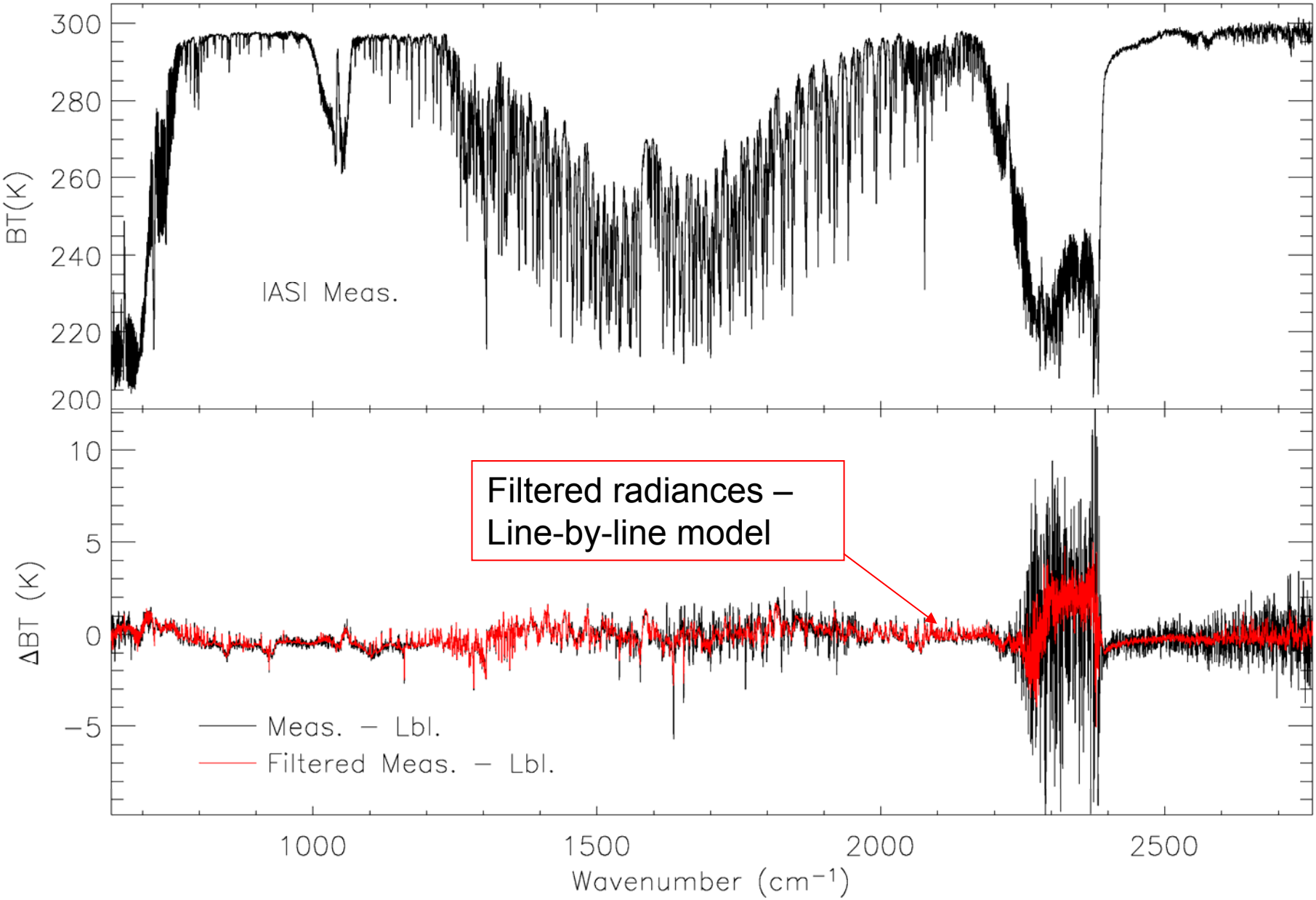
- Eigenvector analysis allows correlated data (*i.e.*, spectral redundancy) to be represented by a relatively small set of functions.
- 8461 channels can easily be represented by a 100 unique coefficients couples with 100 static structure functions (100 x 8461)
- Benefits: Noise filtering and data compression. Distribute and archive 100 coefficients instead of 8461 channels (lossy compression).
- We can now use shortwave IR window channels for applications (e.g., LW versus SW cloud tests)

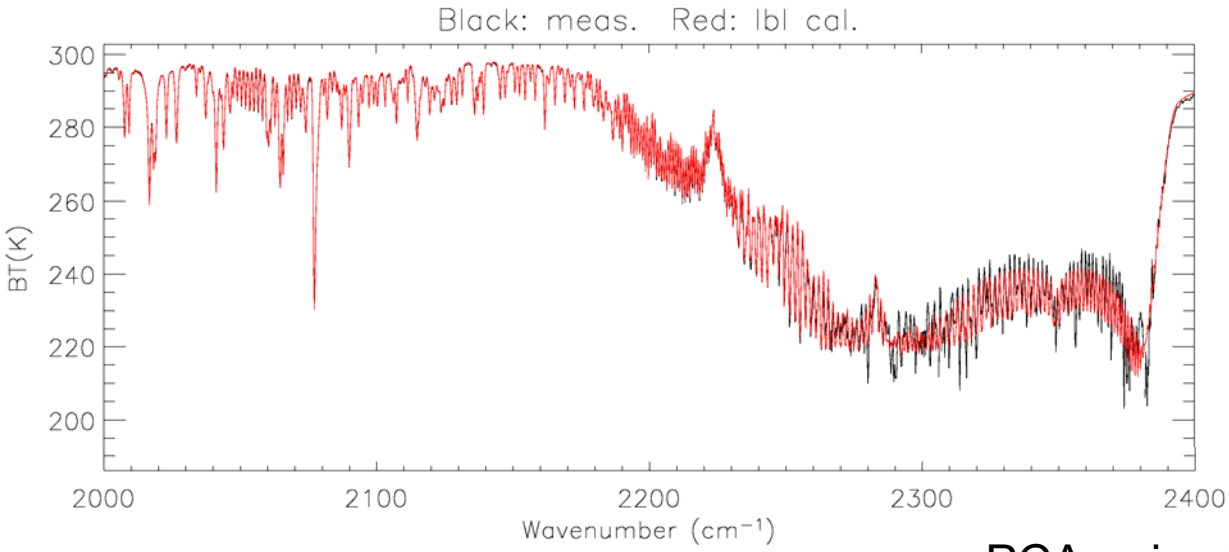
NEDT at 280 K



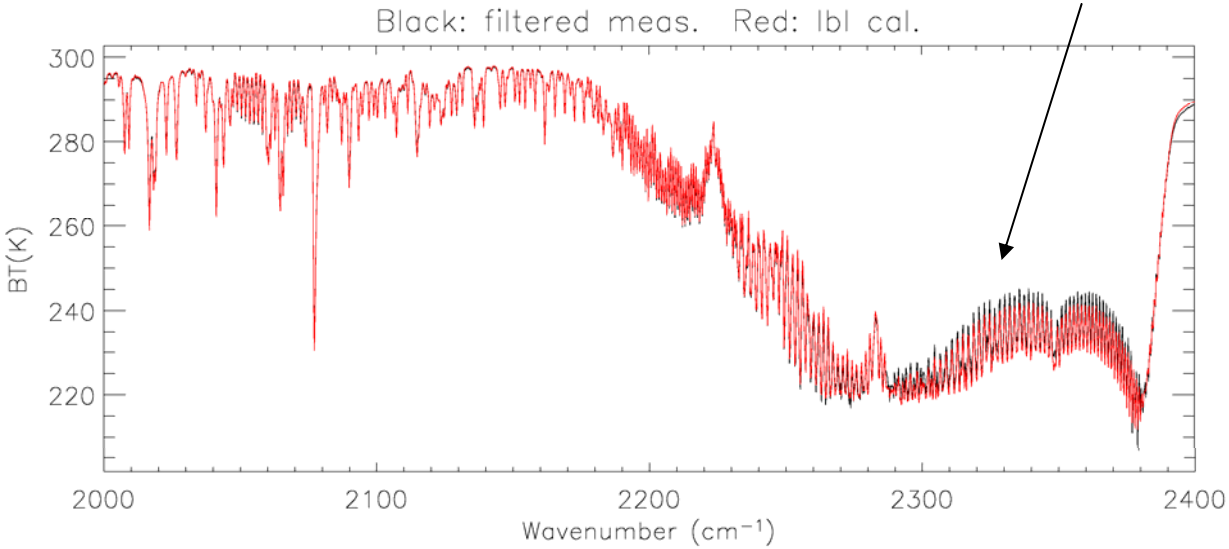
Independent assessment of noise (using Earth scenes) from root mean square difference between measured and reconstructed noise. The reconstructed radiances are noise filtered, therefore the RMS (Obs-Reconstruction) matches the instrument noise.

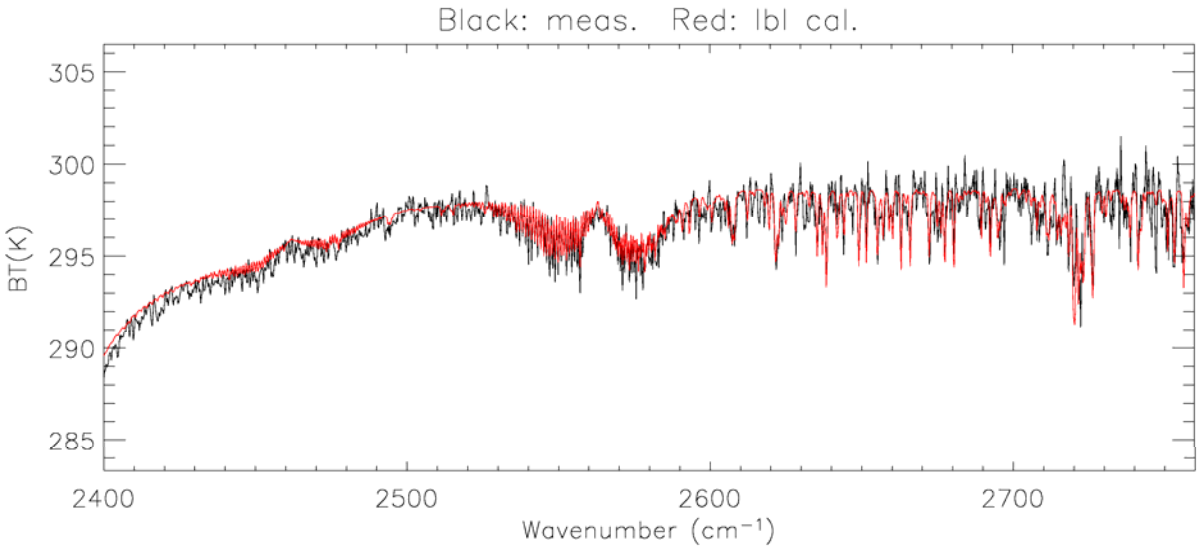
Reconstructed (filtered) radiances can significantly remove random noise from individual spectra.

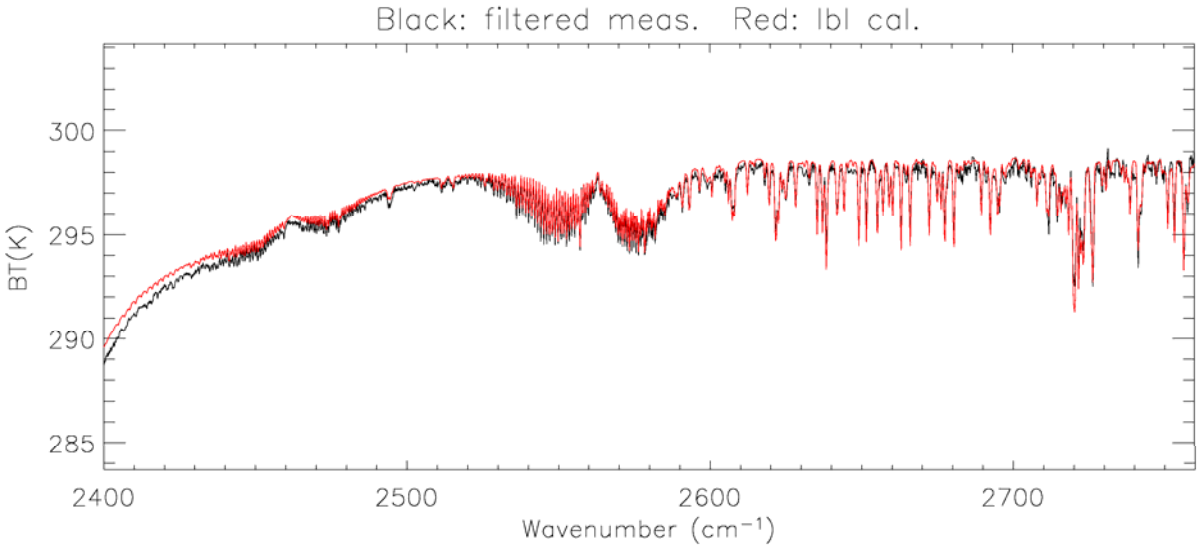




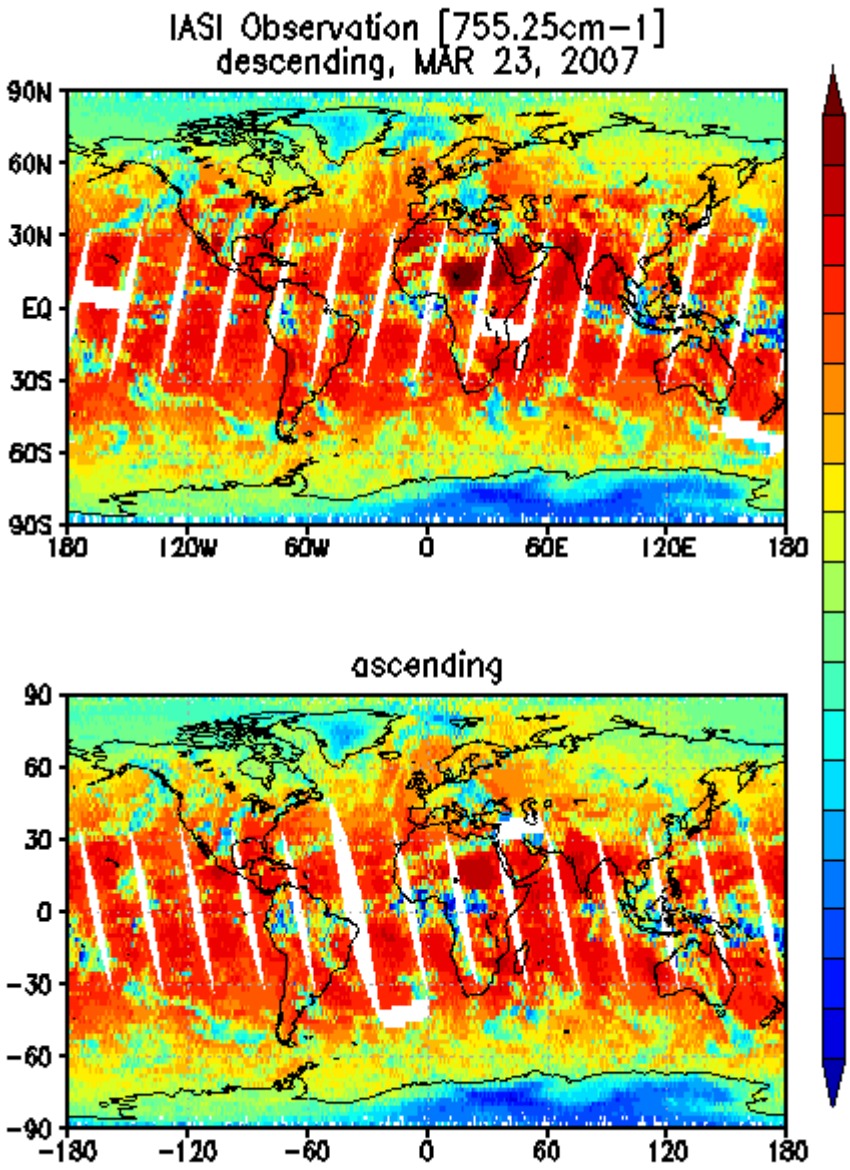
PCA noise filtering

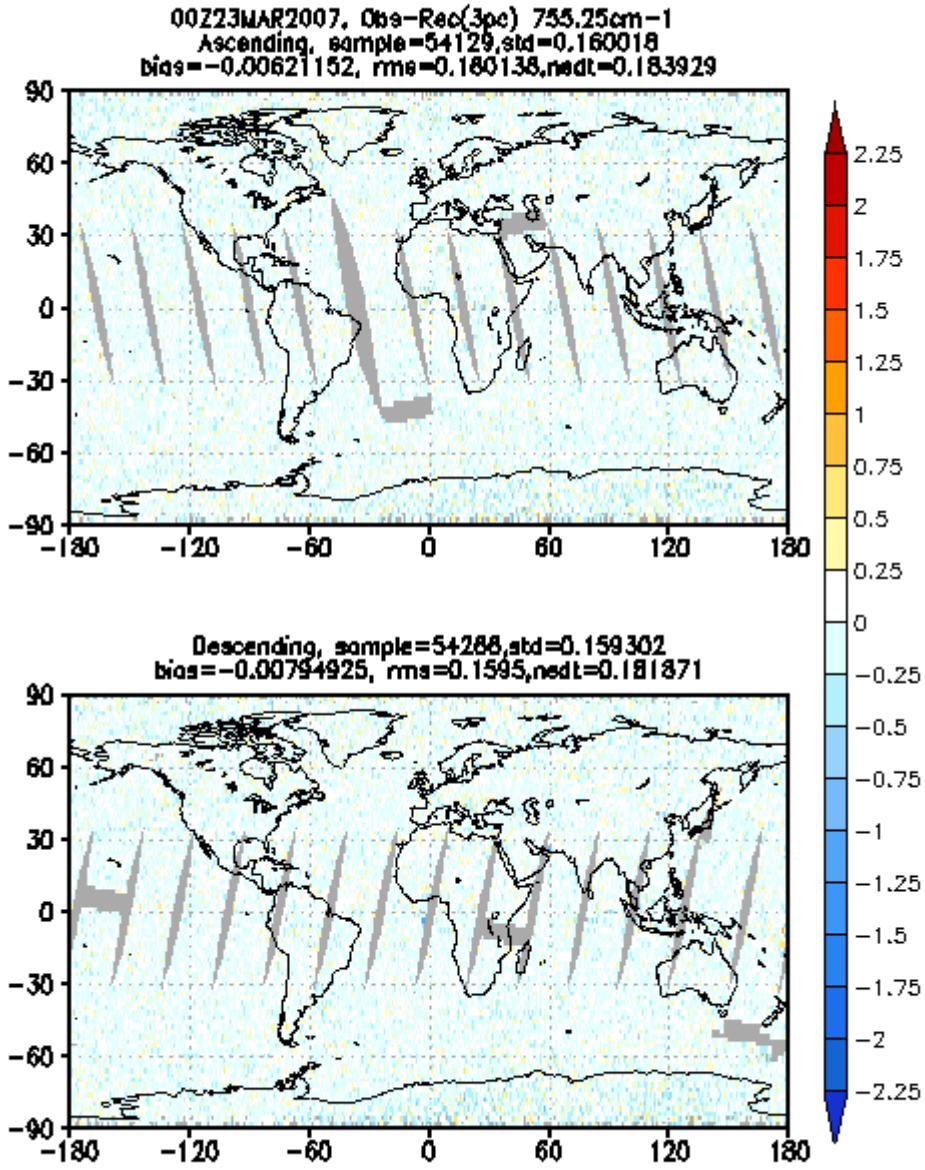



Black: meas. Red: lbl cal.
305
300
295
290
285
BT(K)
2400
2500
2600
2700
Wavenumber (cm^{-1})

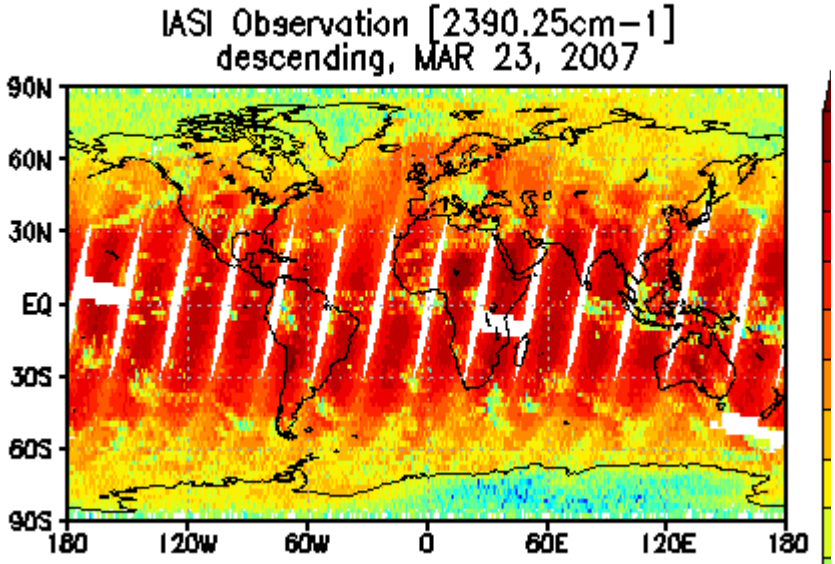

Black: filtered meas. Red: lbl cal.
300
295
290
285
BT(K)
2400
2500
2600
2700
Wavenumber (cm^{-1})

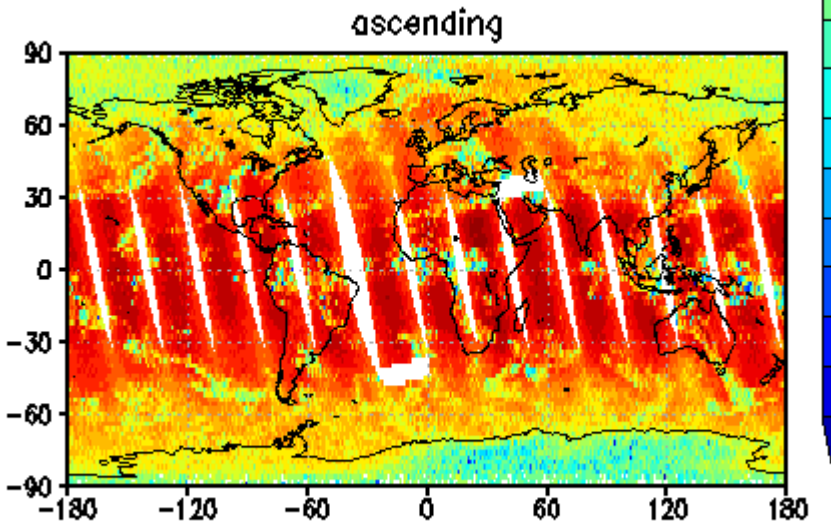
Obs – Reconstruction point by point differences

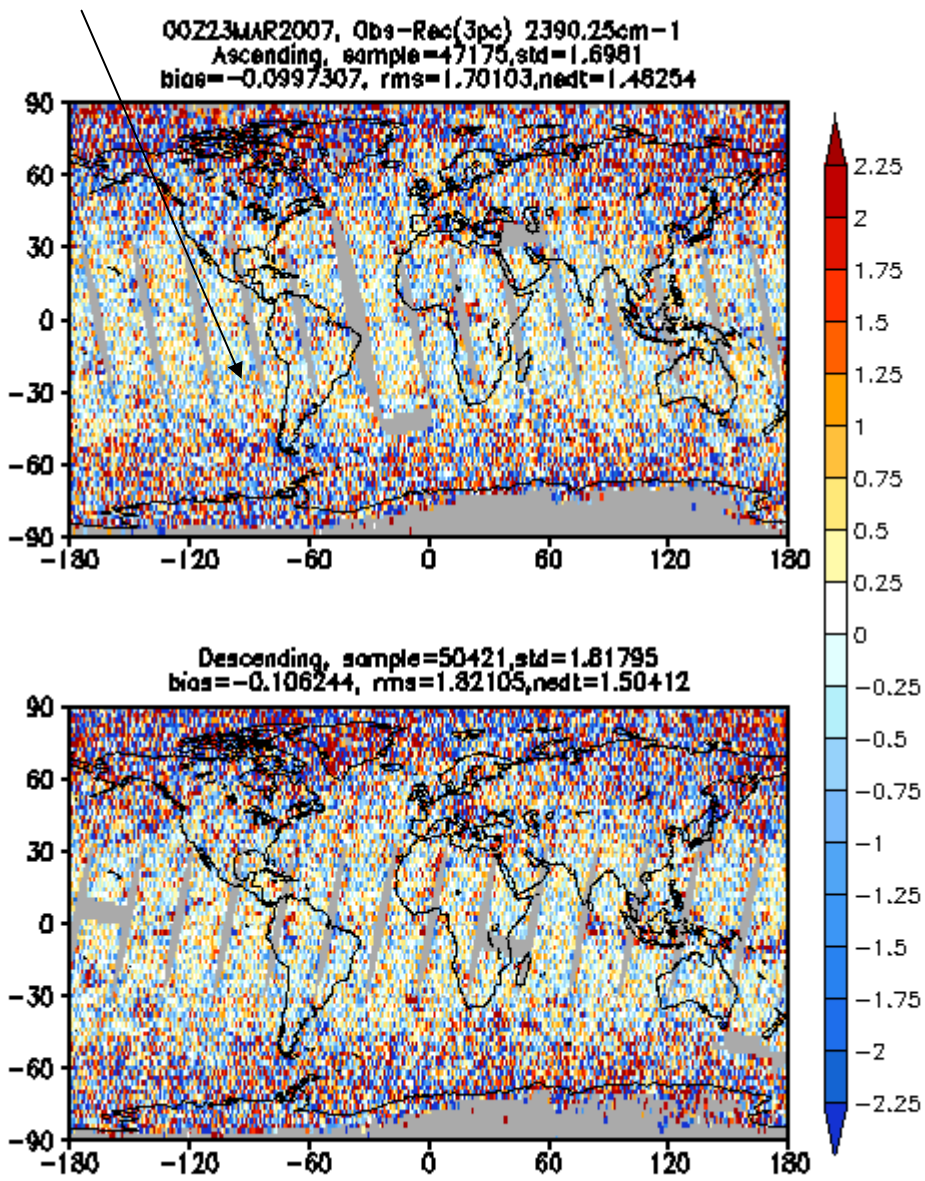




Low values for warm scenes





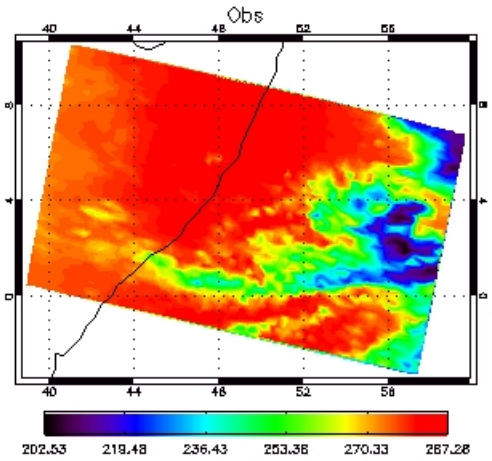


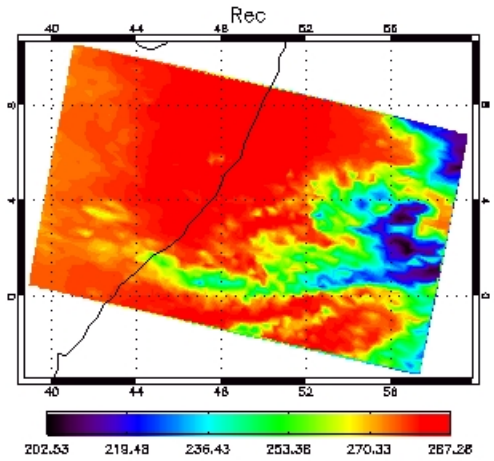
Radiance assimilation should use radiances and not brightness temperatures

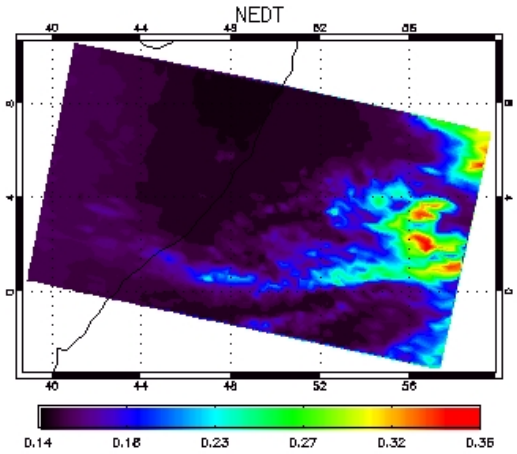
Now we are going to convince you that global (*i.e.*, Static) eigenvectors work!!

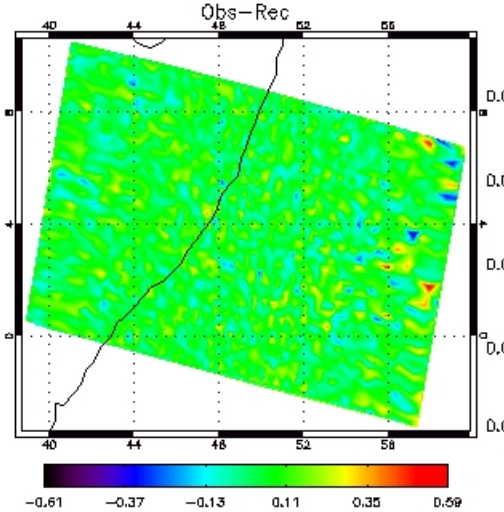
Next few slides show reconstructed versus observed radiances for a "channel" for a given granule in September 2007 using eigenvectors generated from August 3, 2007

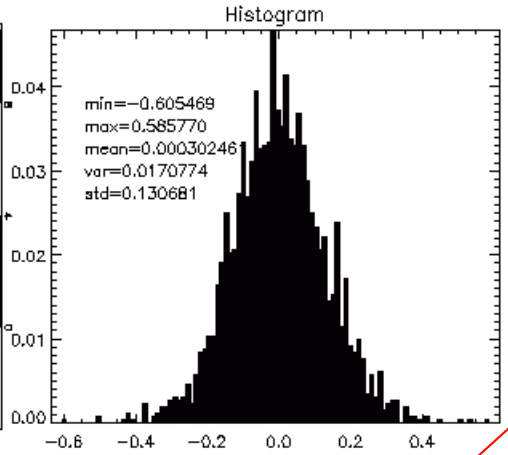
An additional slides (other channels) are at the end of the talk.

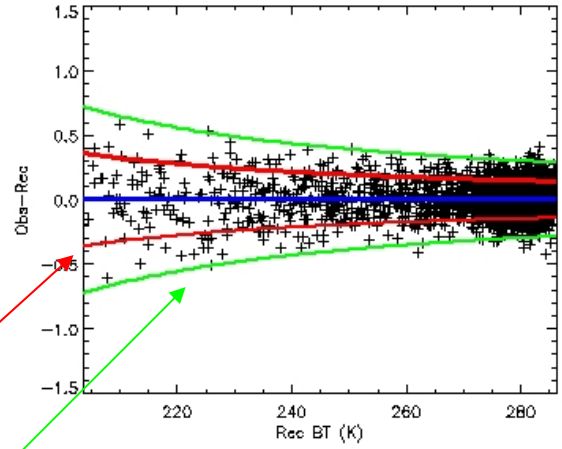








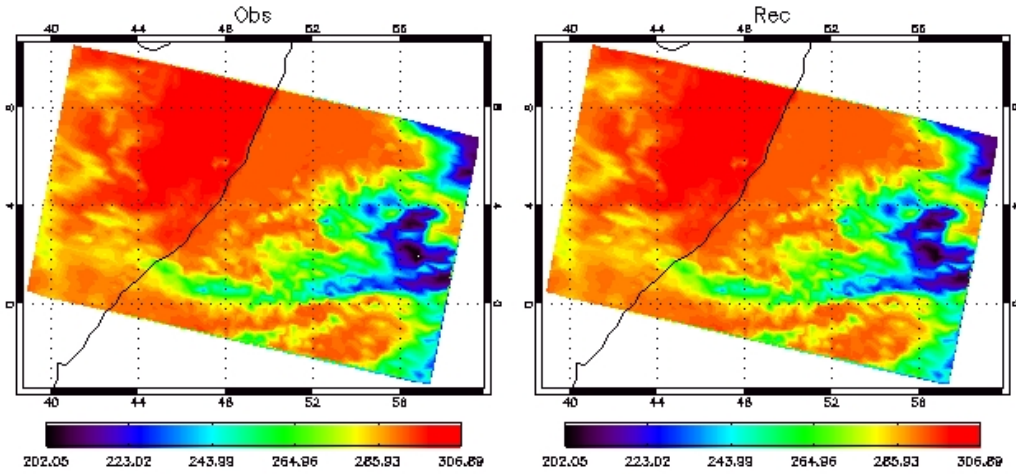


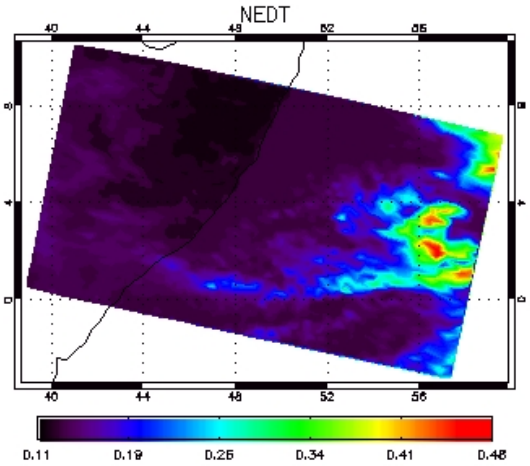


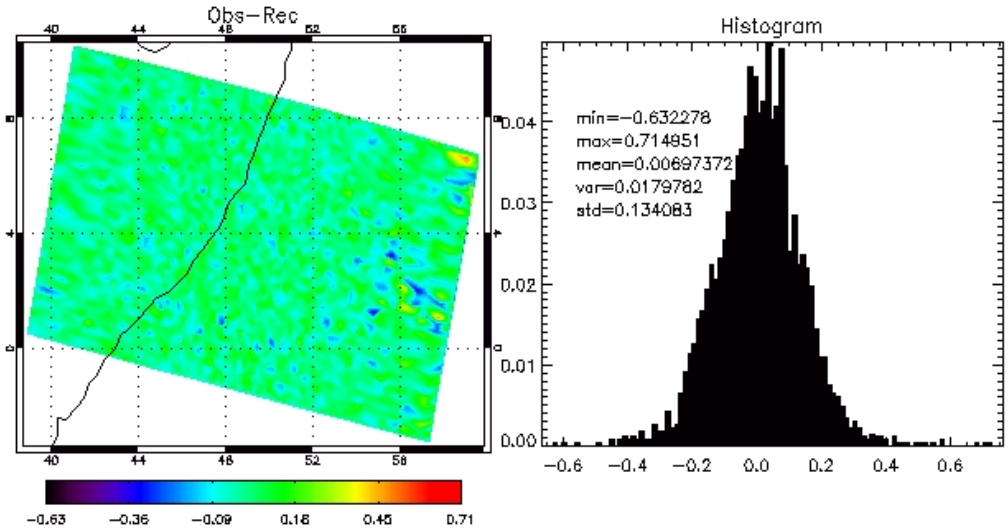
IASI [770.000cm−1] Granule−125 20070910

IASI [770.000cm−1] Granule−125 20070910

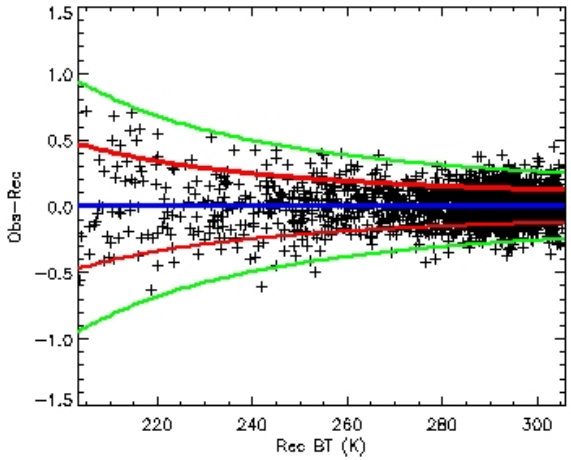
One and two sigma expected NEDT



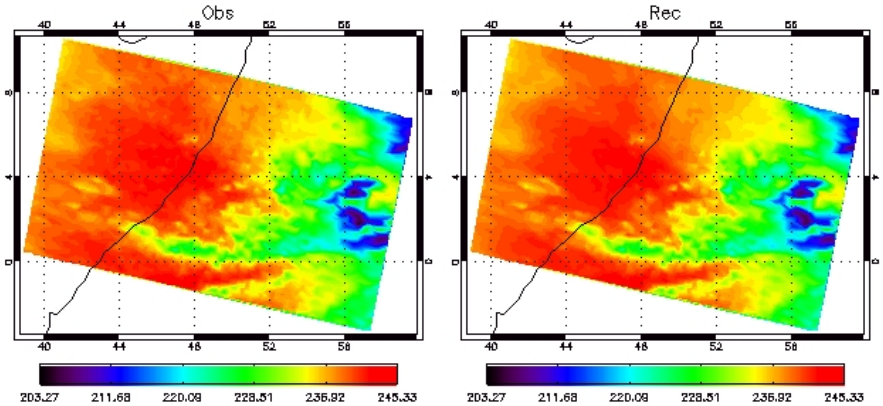


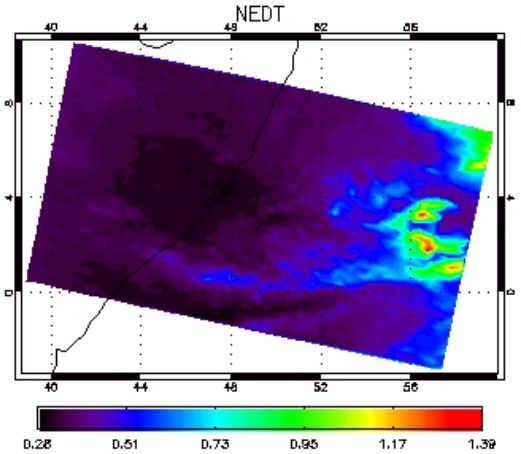


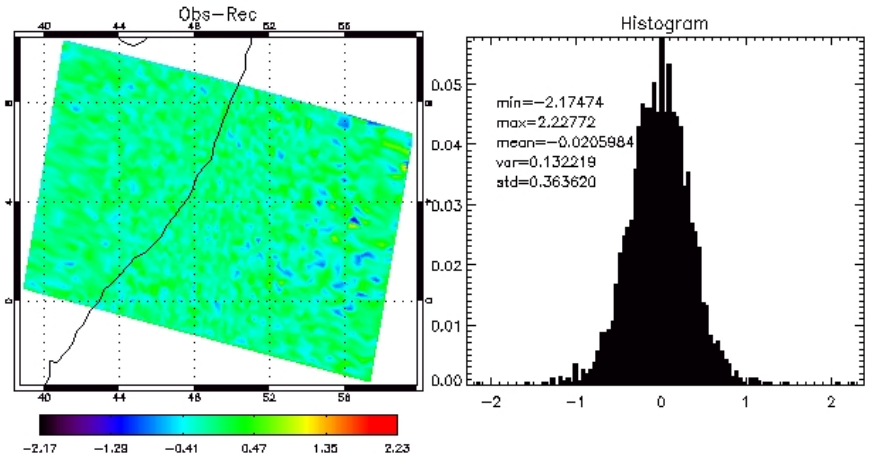
IASI [900.000cm−1] Granule−125 20070910



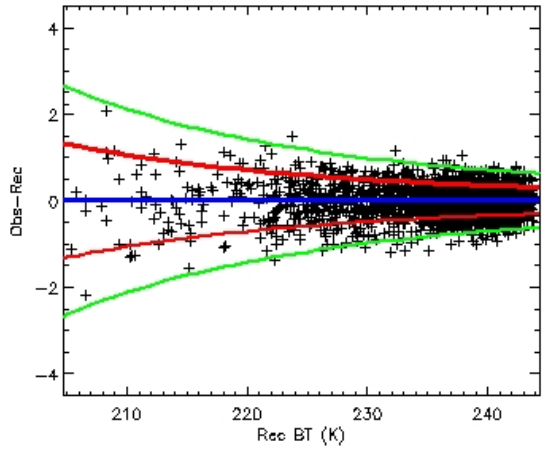
IASI [900.000cm−1] Granule−125 20070910



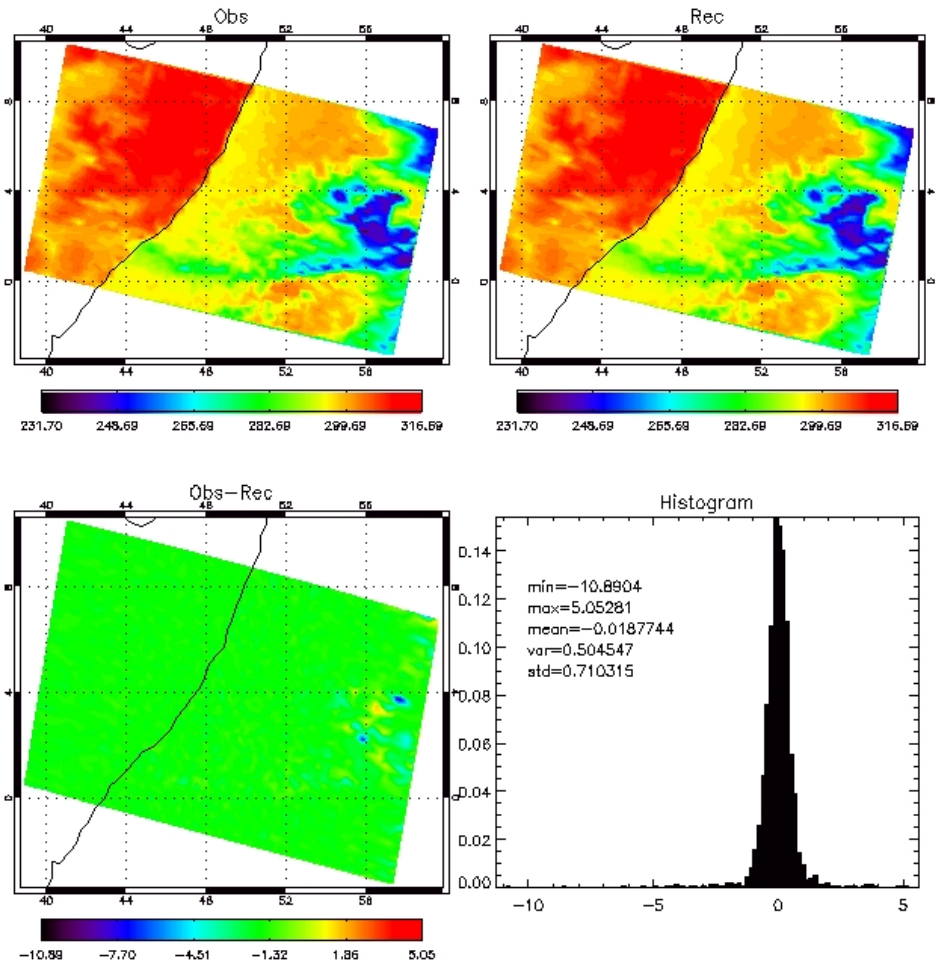




IASI [1594.75cm−1] Granule−125 20070910



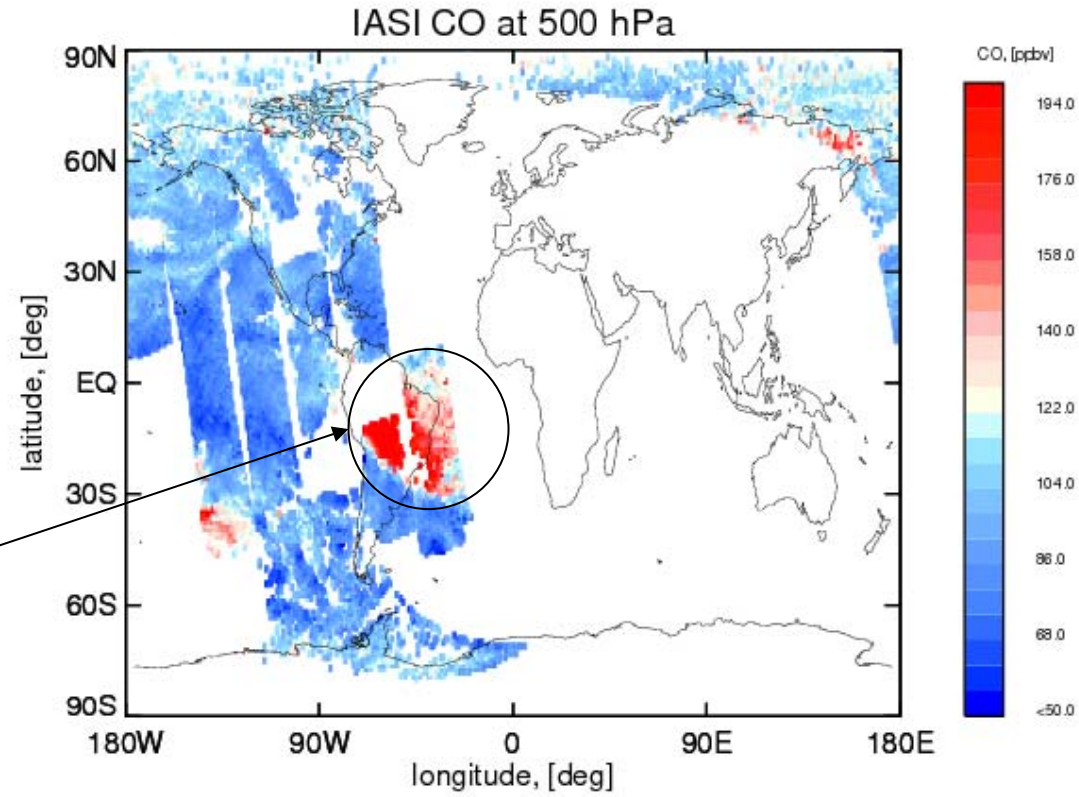
IASI [1594.75cm−1] Granule−125 20070910



IASI [2500cm−1] Granule−125 20070910

IASI [2500cm−1] Granule−125 20070910

IASI CO retrieval at 500 hPa (will be discussed Thursday) for Oct. 19, 2007

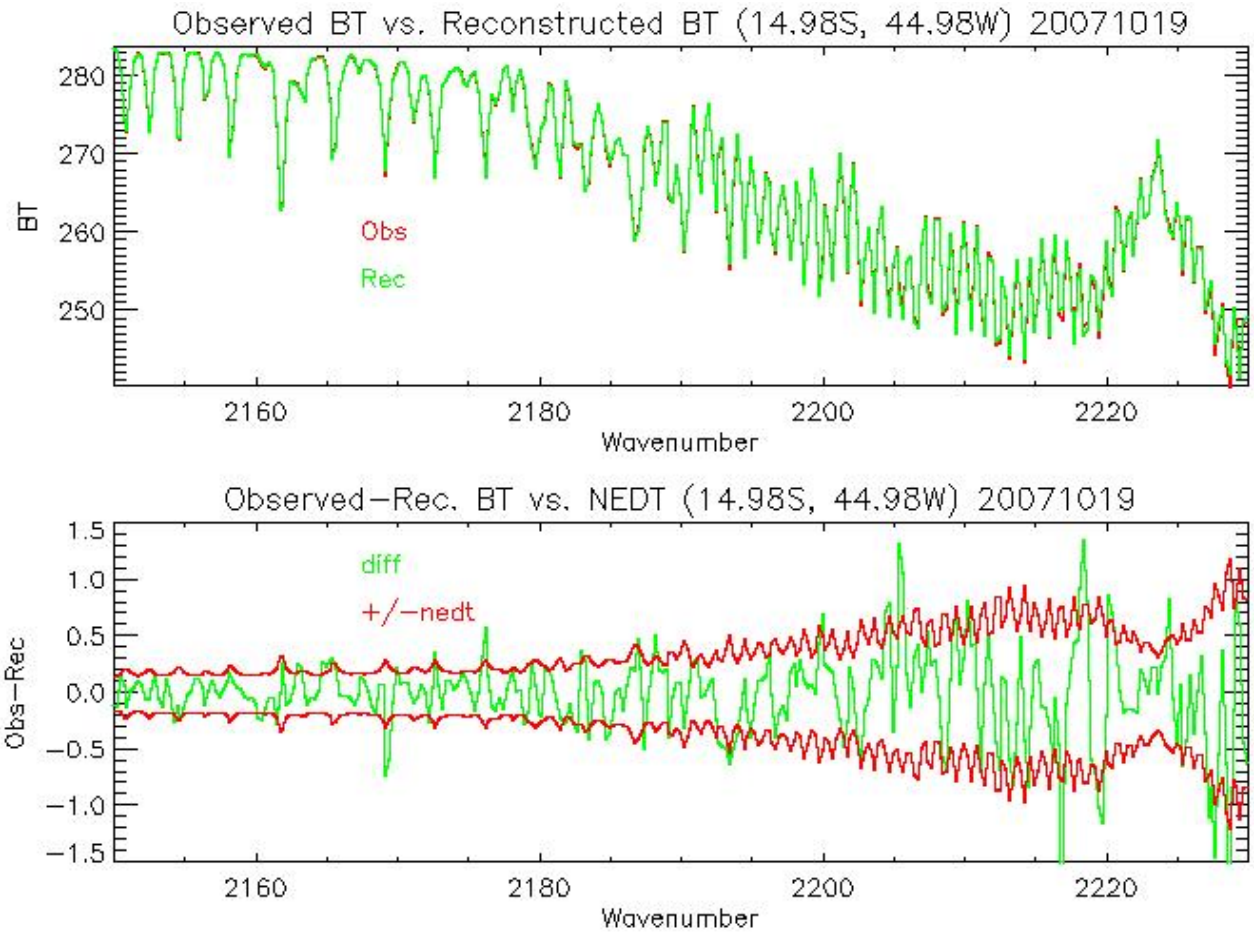


Strong CO

BT Spectrum at 14.98S, 44.98W

Using global eigenvectors

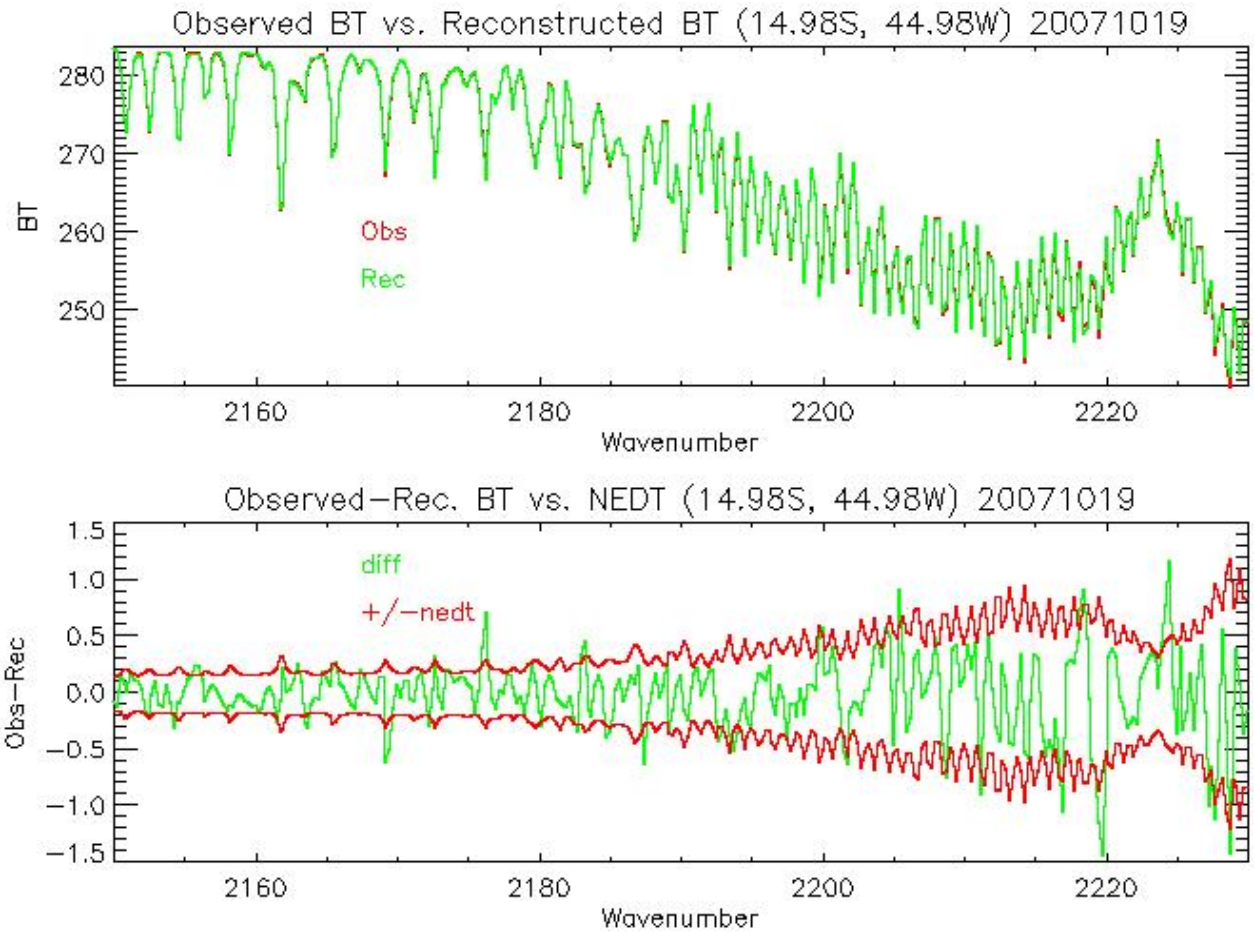
2150 cm^{-1}- 2230 cm^{-1}



BT Spectrum at 14.98S, 44.98W

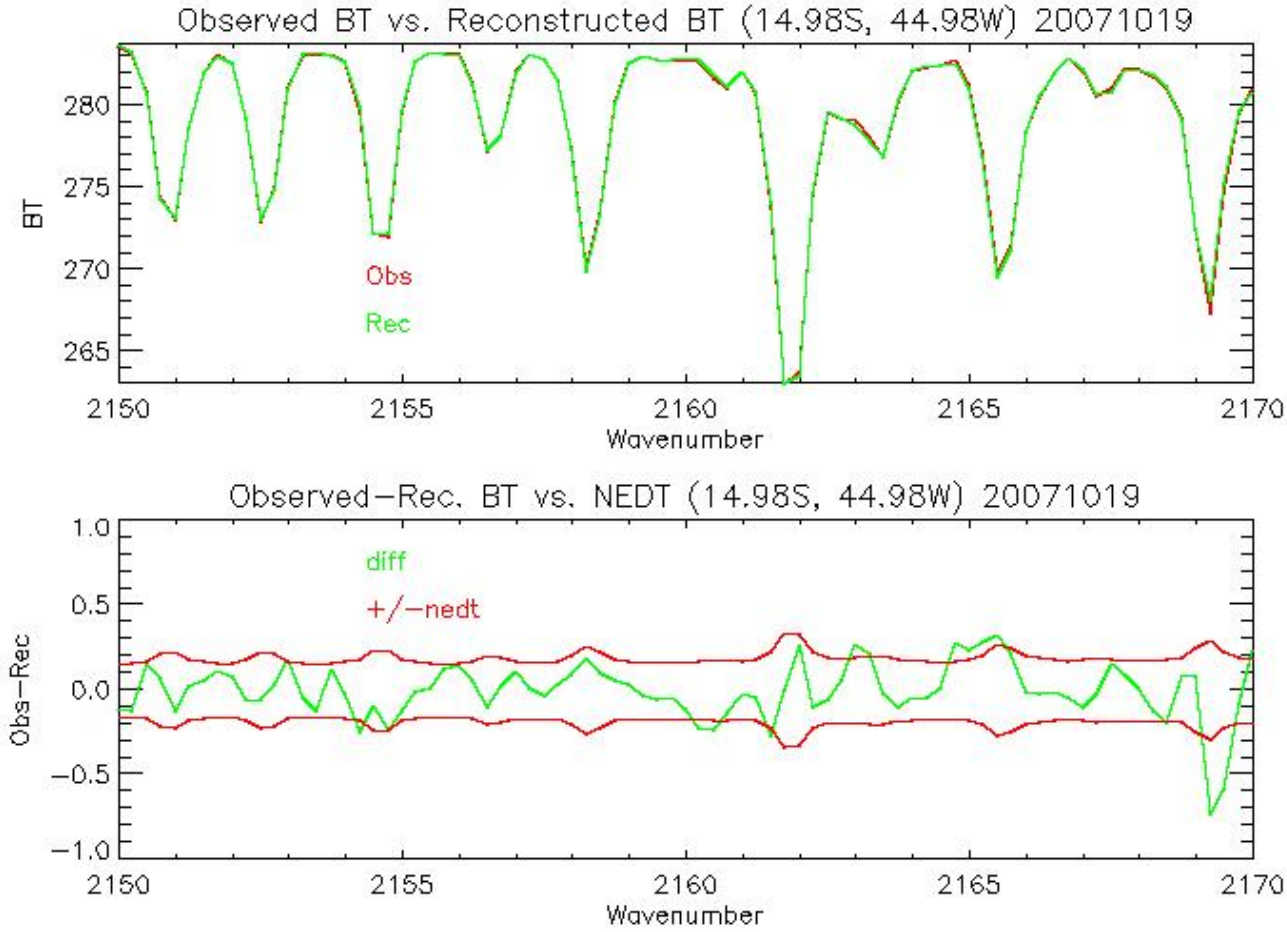
Using granule eigenvector

2150 cm^{-1}- 2230 cm^{-1}



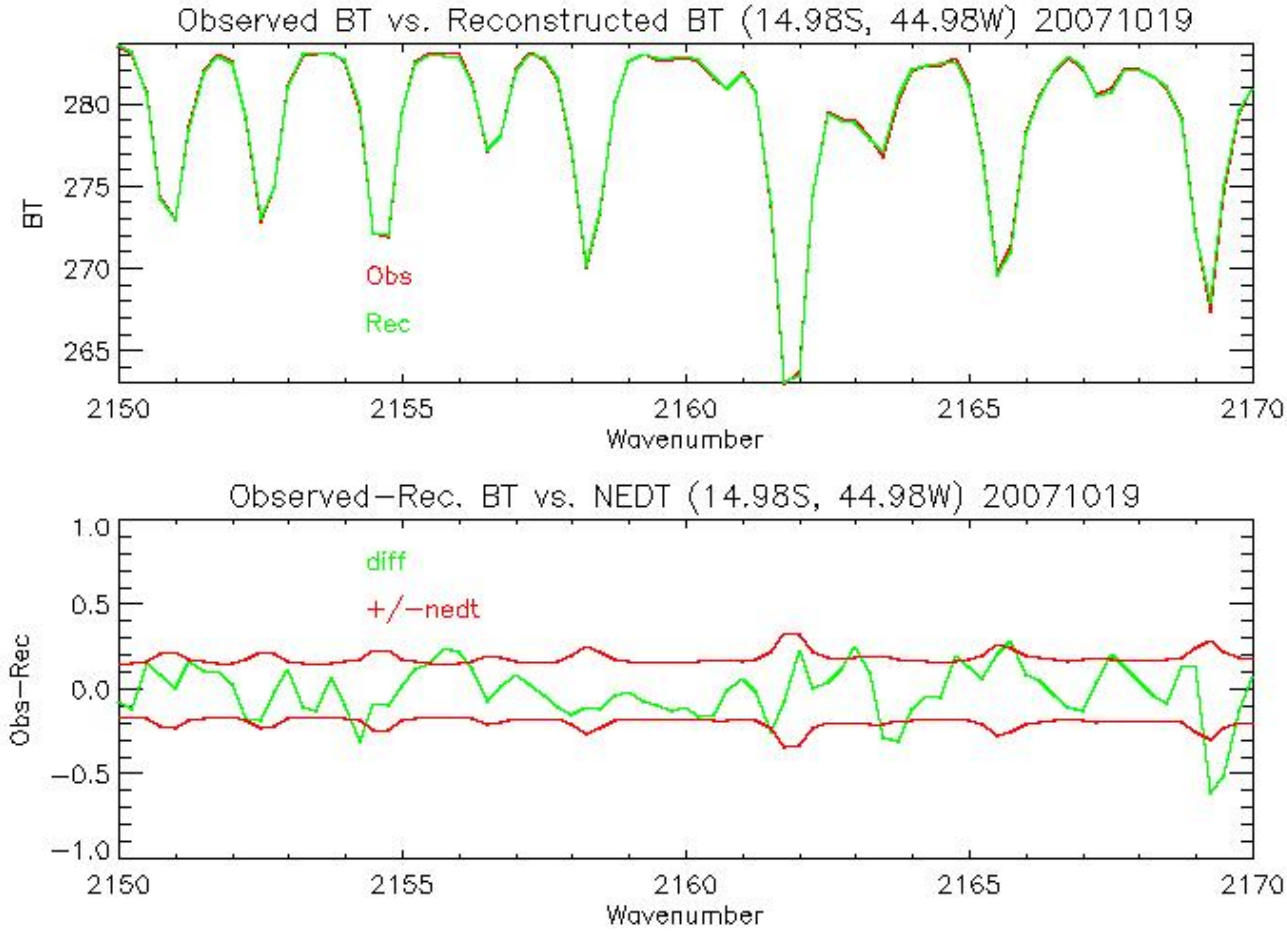
BT Spectrum at 14.98S, 44.98W

Global eigenvector 2150 cm^{-1}- 2170 cm^{-1}



BT Spectrum at 14.98S, 44.98W

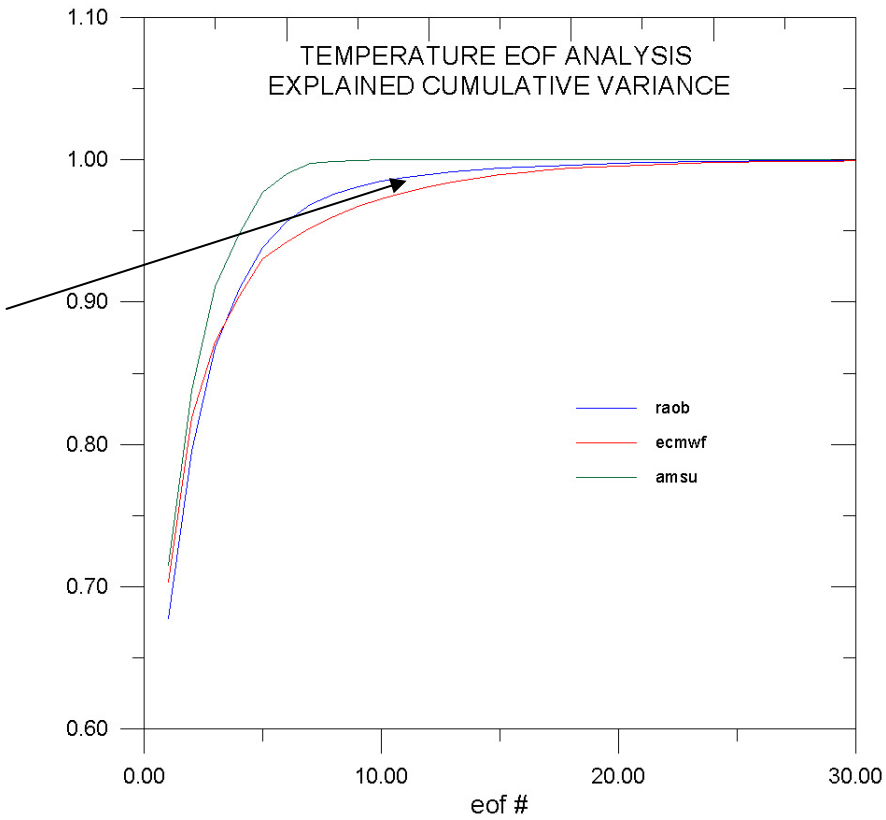
Granule eigenvector 2150 cm^{-1}- 2170 cm^{-1}



IASI Regression Retrieval—for Temperature and Moisture Profiles

Regression Retrieval Training

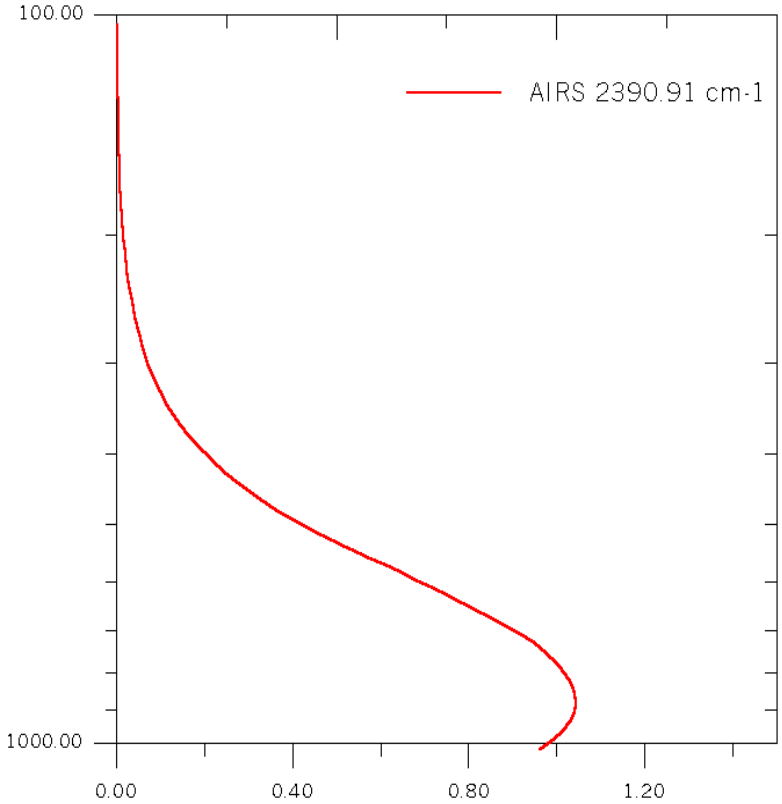
- Train PCS's against ECMWF analysis
  - We have confirmed over the past 5 years with AIRS that ECMWF represents the true atmosphere extremely well.
  - ECMWF has a similar number of independent pieces of information as RAOB's
- Our regression training set is currently for May 4, 2007
  - NOTE: Eigenvectors are trained on Aug. 4, 2007.
- The training pairs consists of ECMWF atmospheric profiles and partly cloudy/clear IASI FOVs

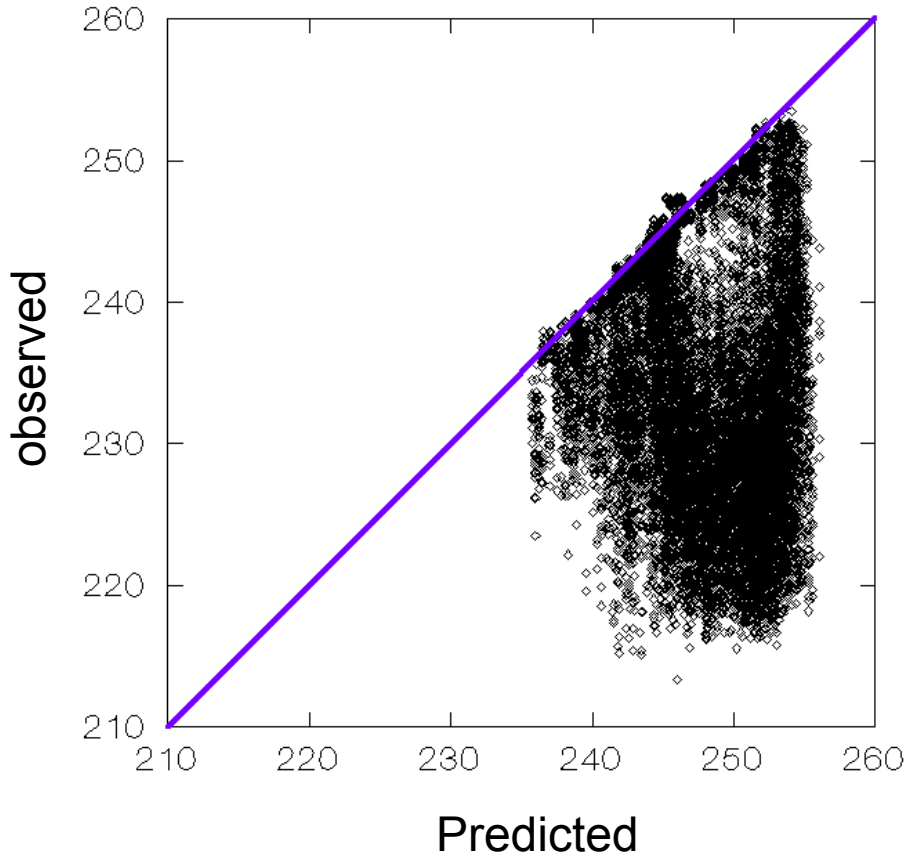


Identification of IASI Partly Cloudy/Clear FOVS

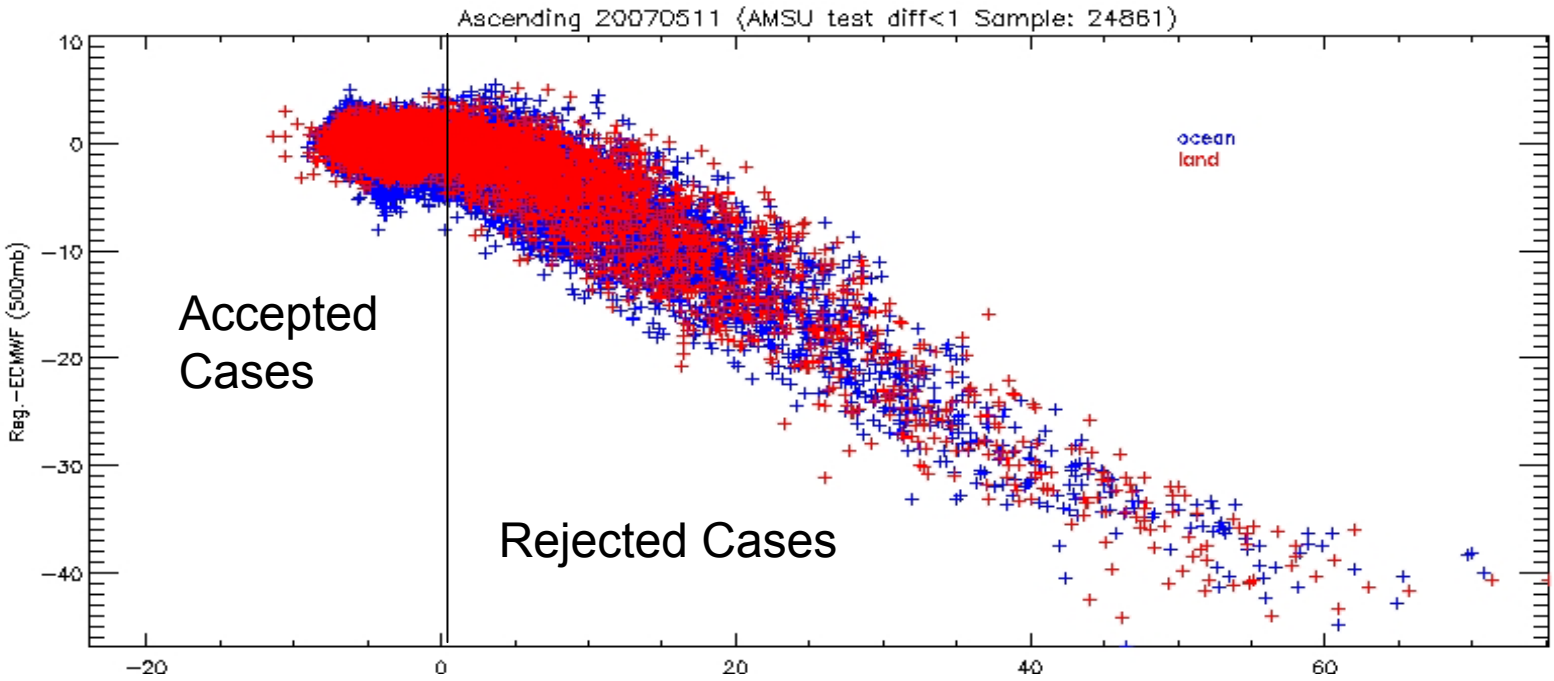
- We predict reconstructed IASI 2390.25 cm^{-1} channel (to de-noise) from AMSU channels 4, 5 and 6.
- If the difference of the observed minus the predicted is less than zero – the FOV is classified as Partly Cloudy/Clear FOV
- This test works very well over ocean and snow/ice free land.
- Also this test cannot detect low clouds.
- However above 700 hPa, the regression retrievals are very accurate, and the global coverage is about 70%!
- Since AIRS temperature retrievals have shown positive impacts using NASA GEOS-5 forecast system, and ECWMF are essentially assimilating band 1 IASI channels peaking from the mid troposphere and above, it may make sense to test these retrievals in NWP

AIRS 2390.91cm^{-1} Weighting function

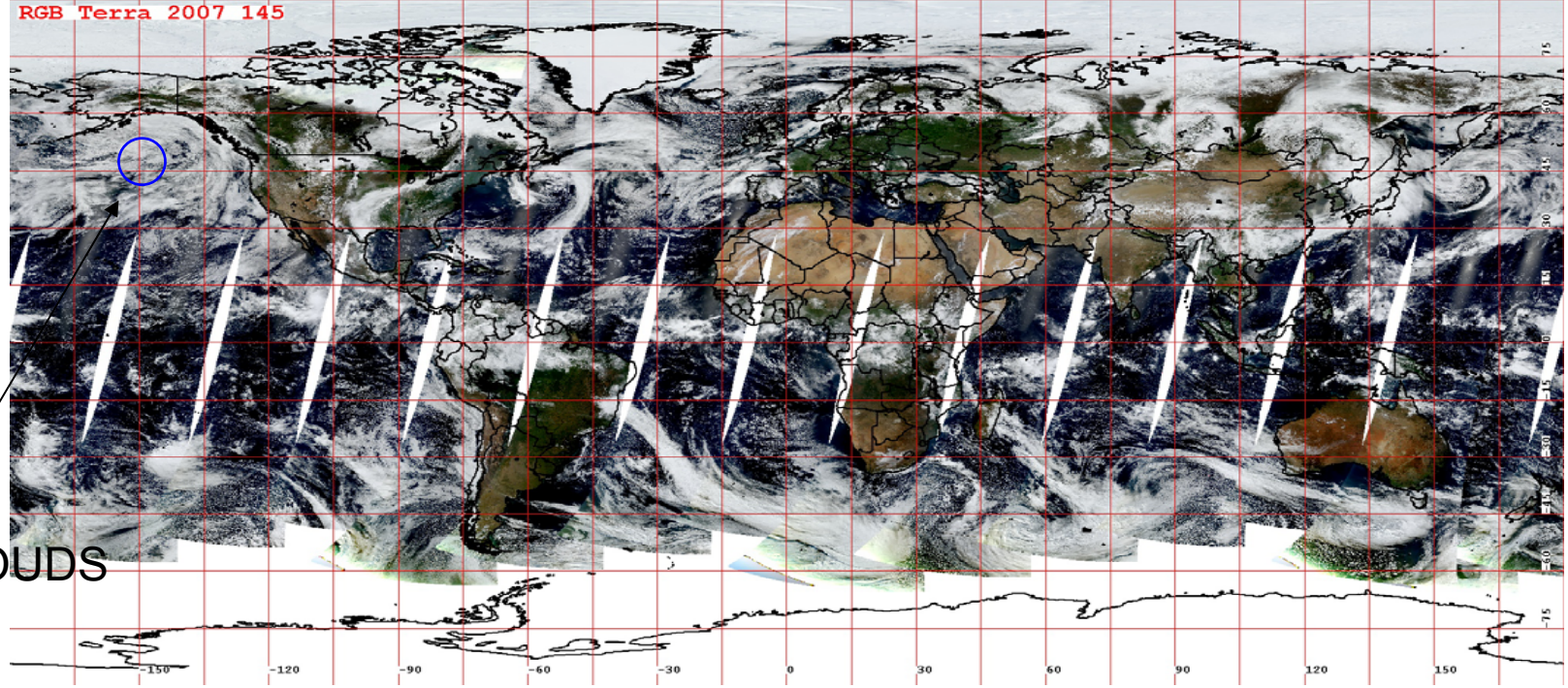




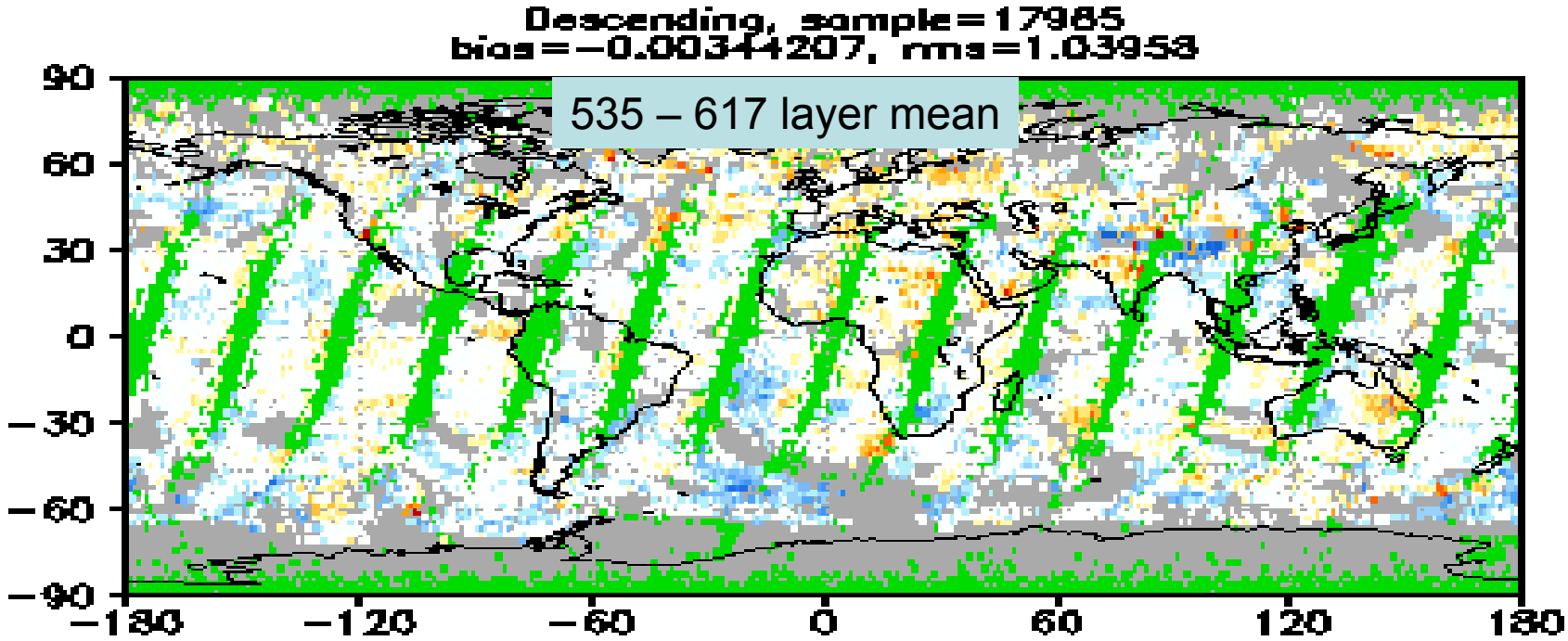
The “2390 cm^{-1} diff test” removes overcast and precipitating clouds



2390.91 cm^{-1} diff test


RGB Terra 2007 145

LOW CLOUDS


Descending, sample=17985
bias=−0.00344207, rms=1.03958
535 – 617 layer mean
90
60
30
0
−30
−60
−90
−180
−120
−60
0
60
120
180

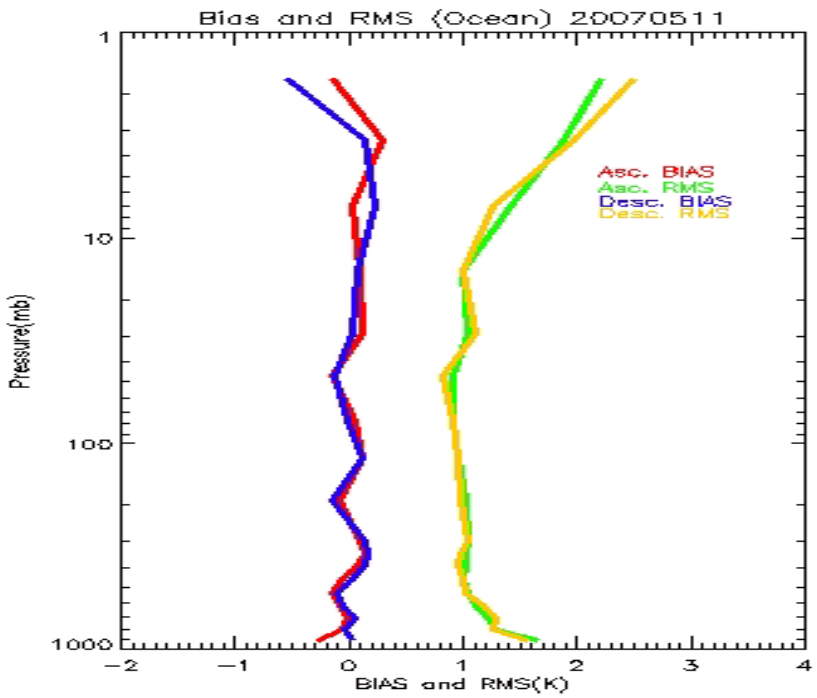
Real-time Steps

- Read IASI PC Scores BUFR file (operational on Oct. 30)
- Reconstruct 2390.25 cm^{-1} brightness temperature
- Predict 2390.25 cm^{-1} from AMSU
- If diff < 0, generate regression retrieval using the first 85 PCS.
- For now just use product above 700 hPa and for surface temperature above 271.5 K

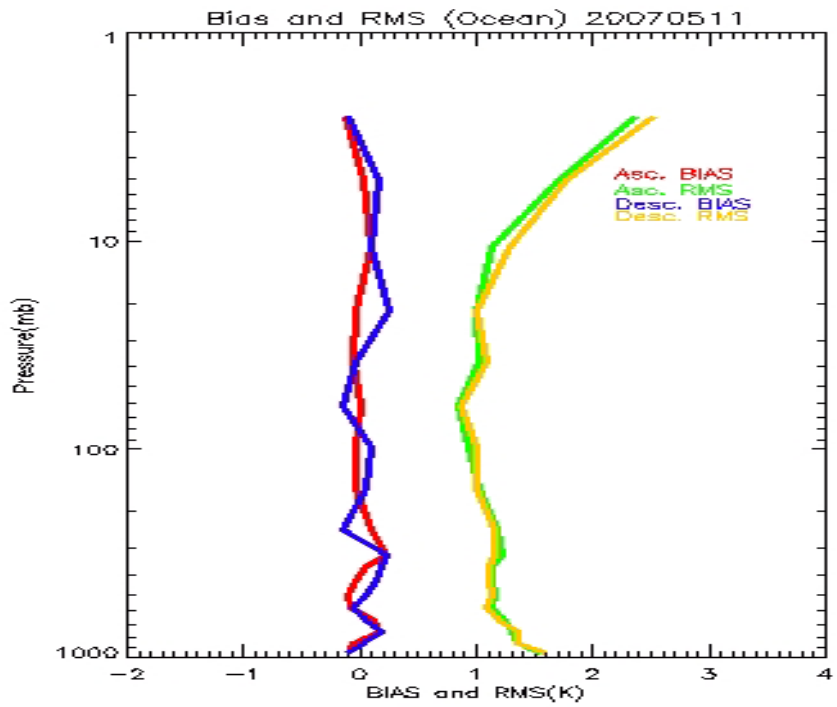
So now for the results on data from May 25, 2007. We compare retrieval bias and rms for AMSU only vs IASI plus AMSU

IASI and AIRS Regression Temperature Retrieval Accuracy

IASI



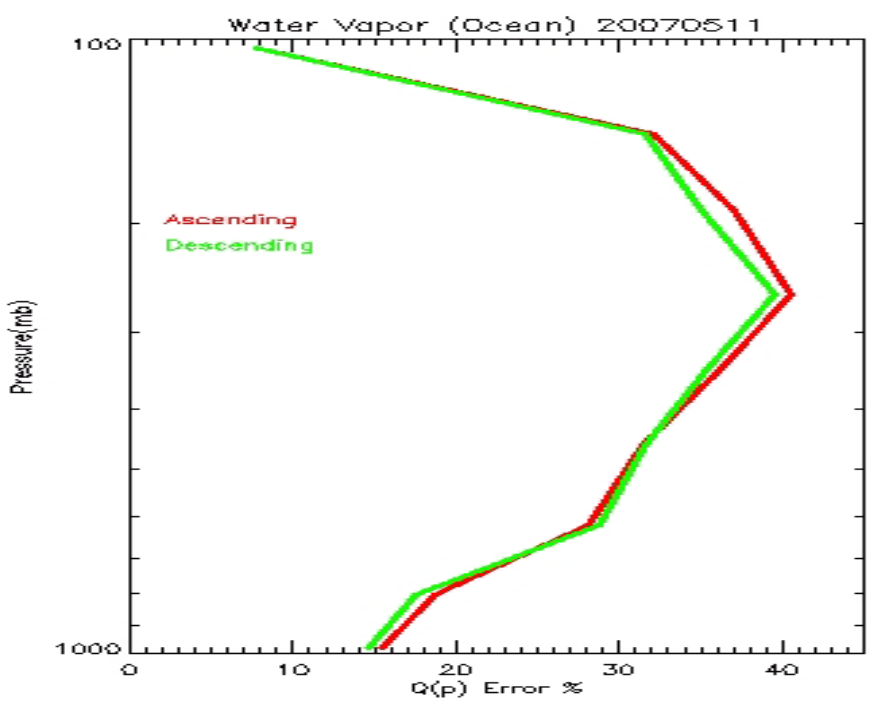
AIRS



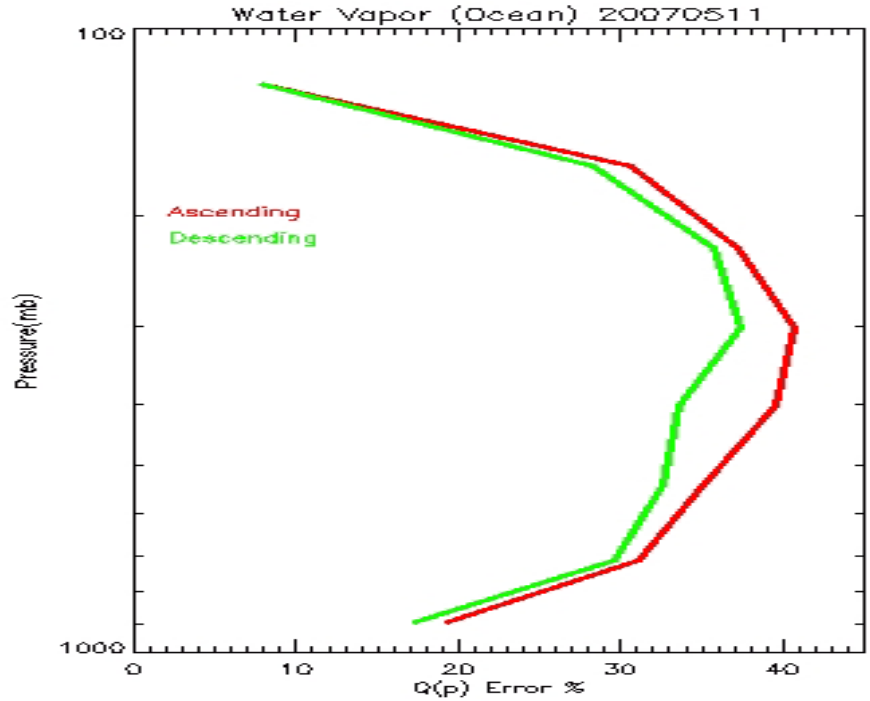
Independent dataset (ocean only)

IASI and AIRS Regression Moisture Retrieval Accuracy

IASI

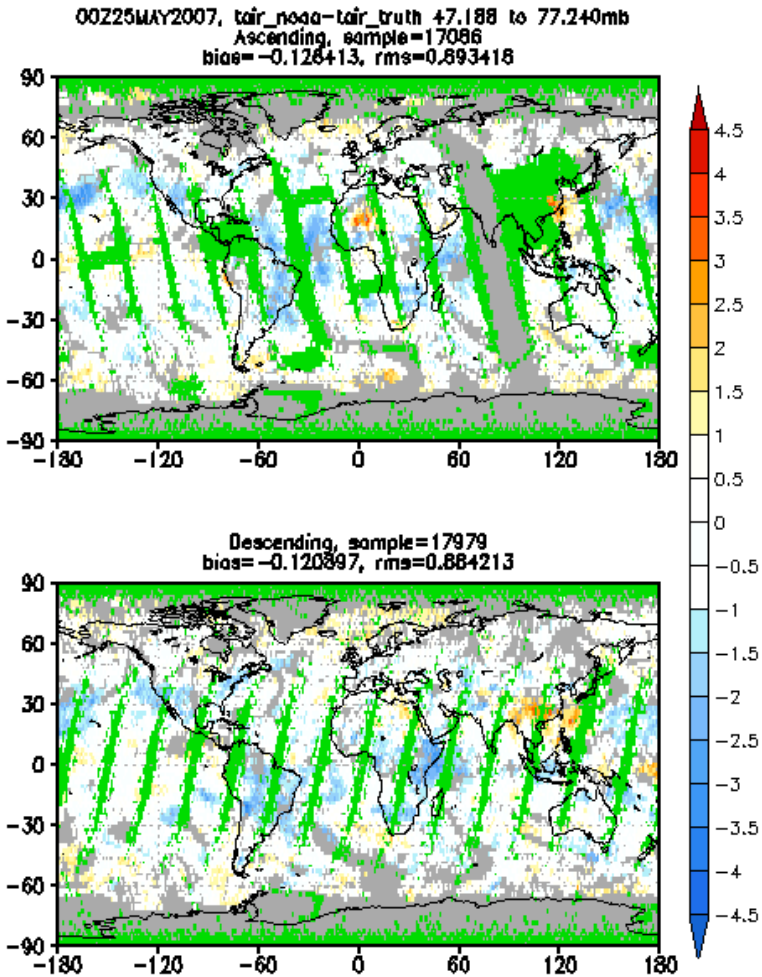


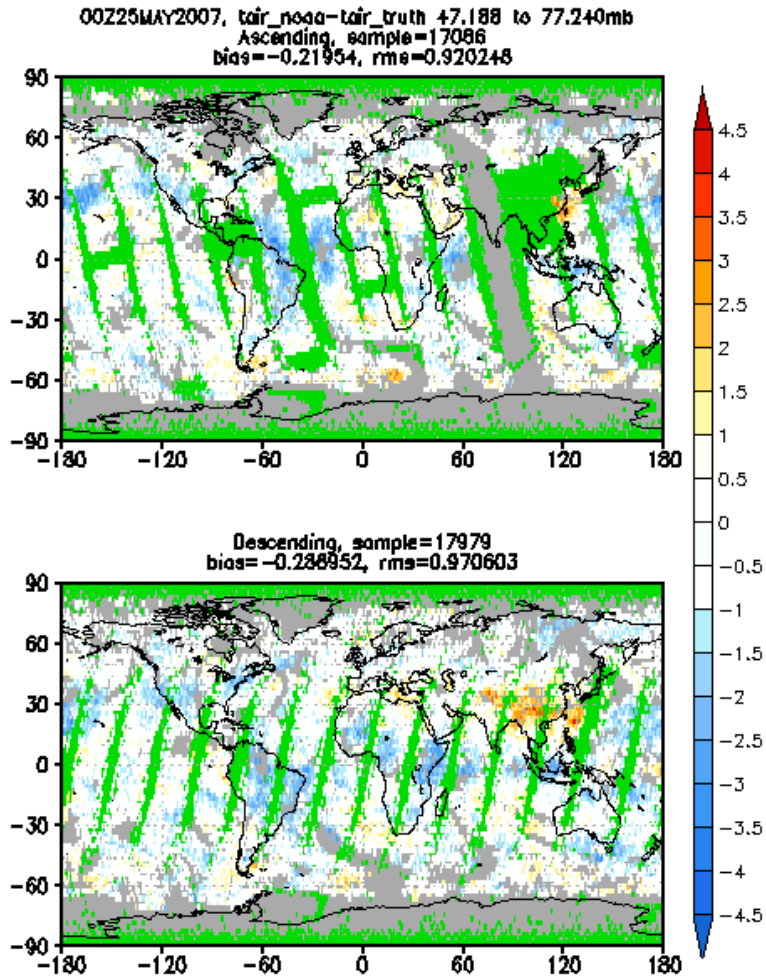
AIRS



Independent dataset (ocean only)

Regression-ECMWF: Temperature at 47.188 — 77.24 hPa

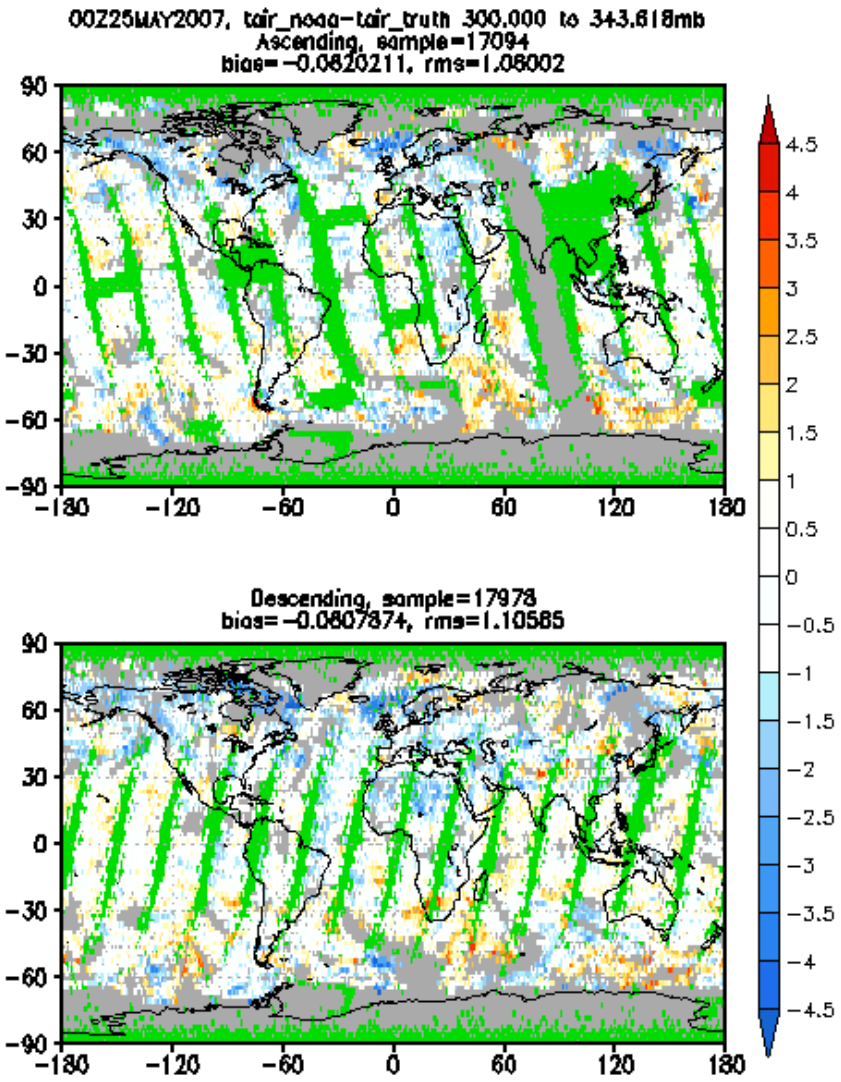




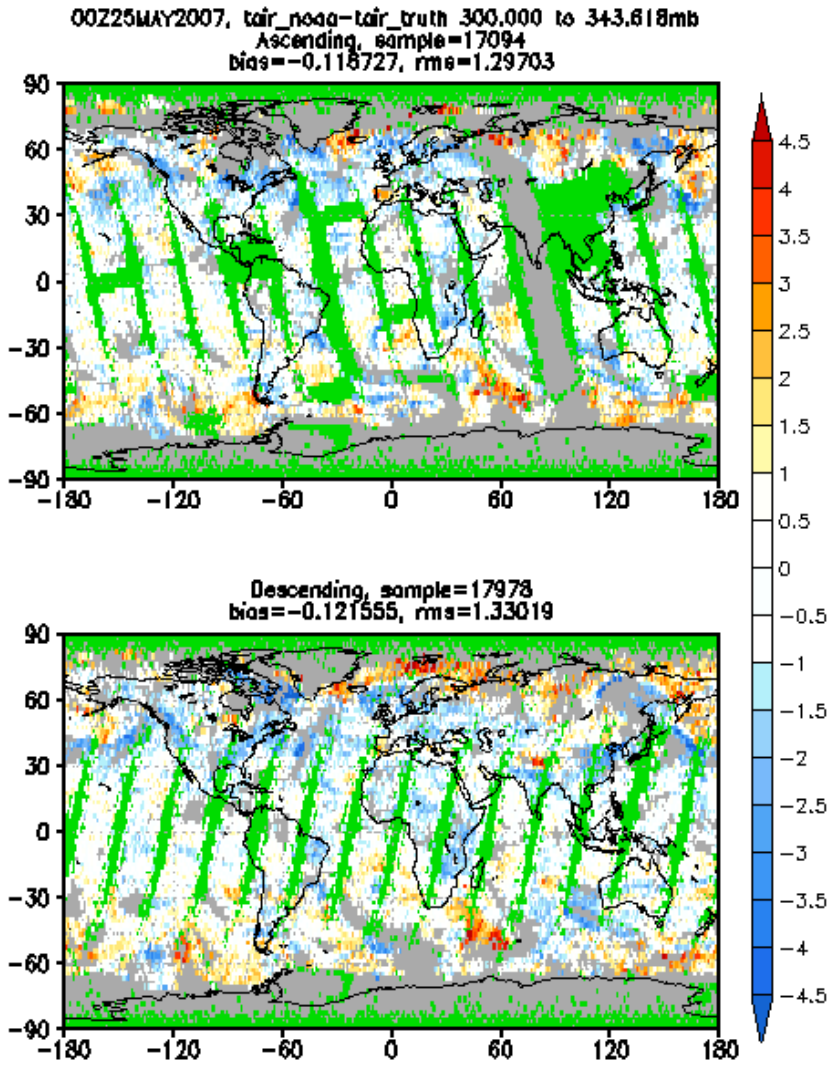
Predictors: AMSU + IASI

Predictors: AMSU-only

Regression-ECMWF: Temperature at 300 — 343.618 hPa

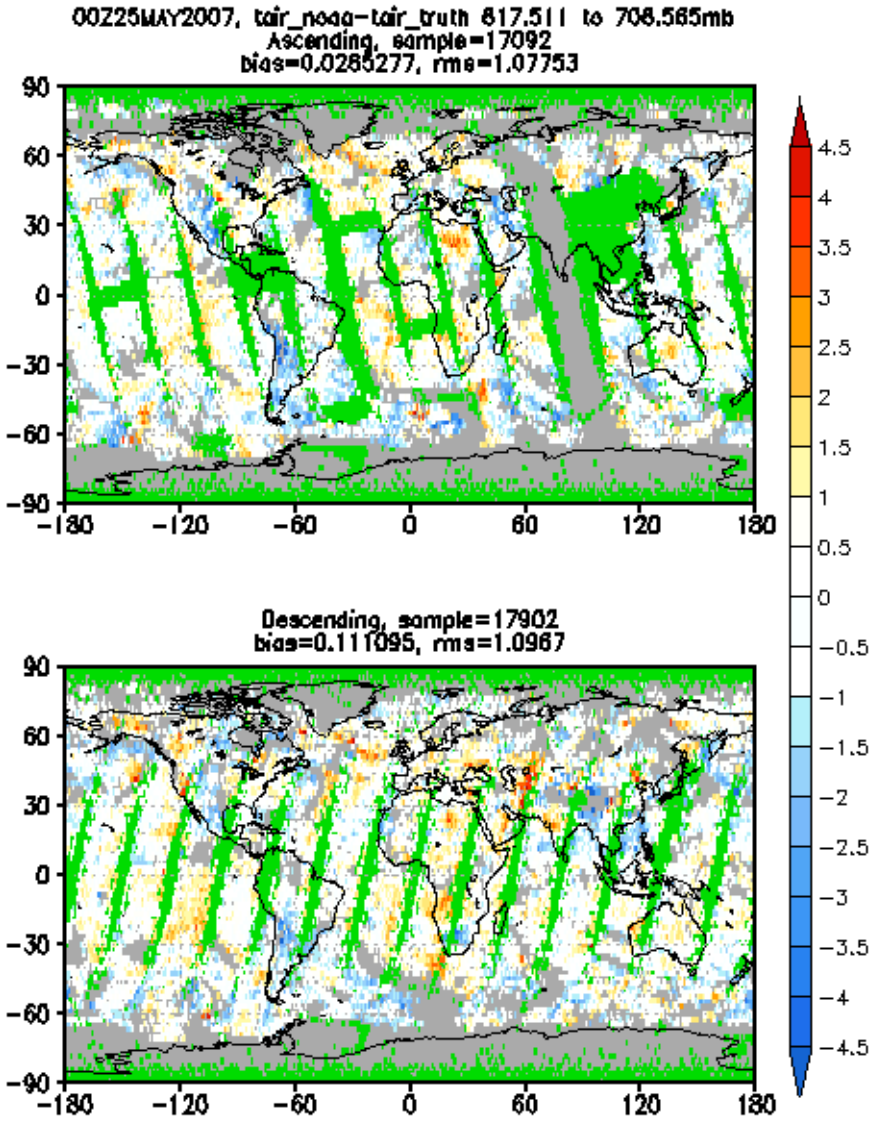


Predictors: AMSU + IASI

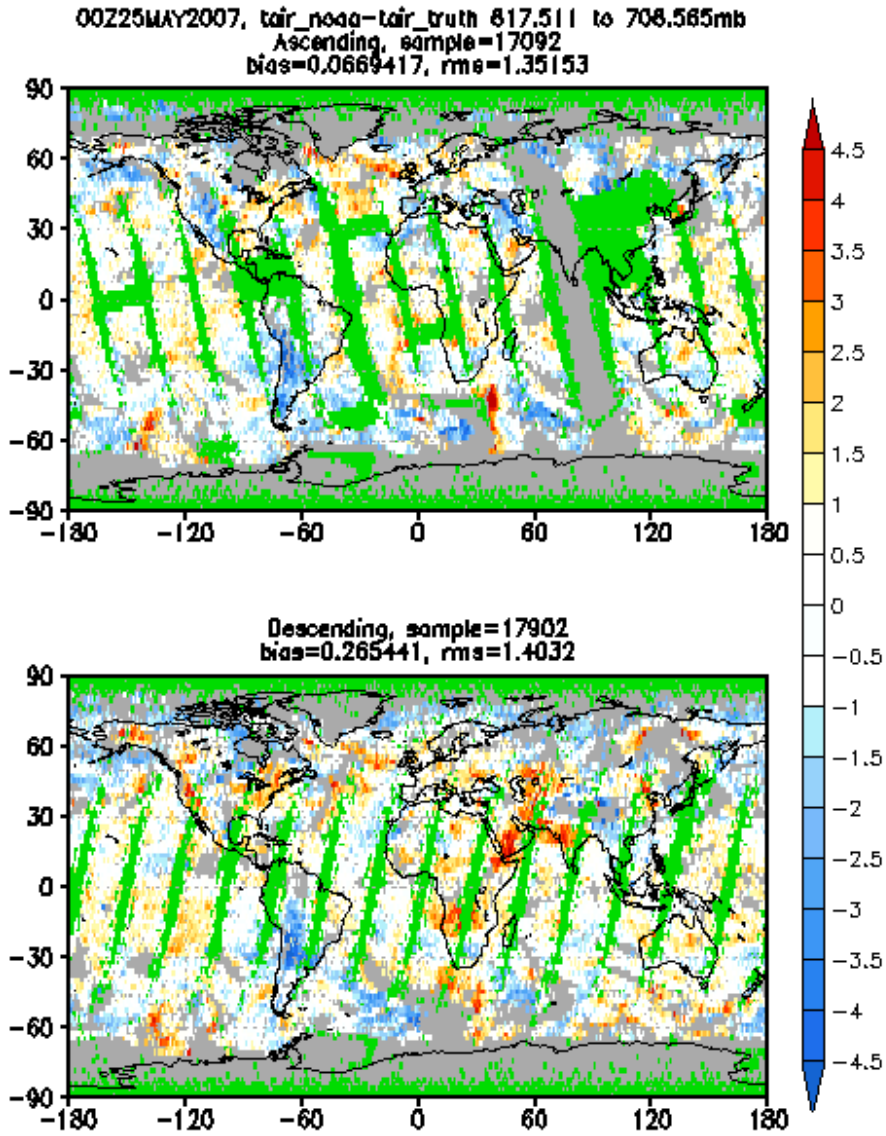


Predictors: AMSU

Regression-ECMWF: Temperature at 617.511 — 706.565 hPa

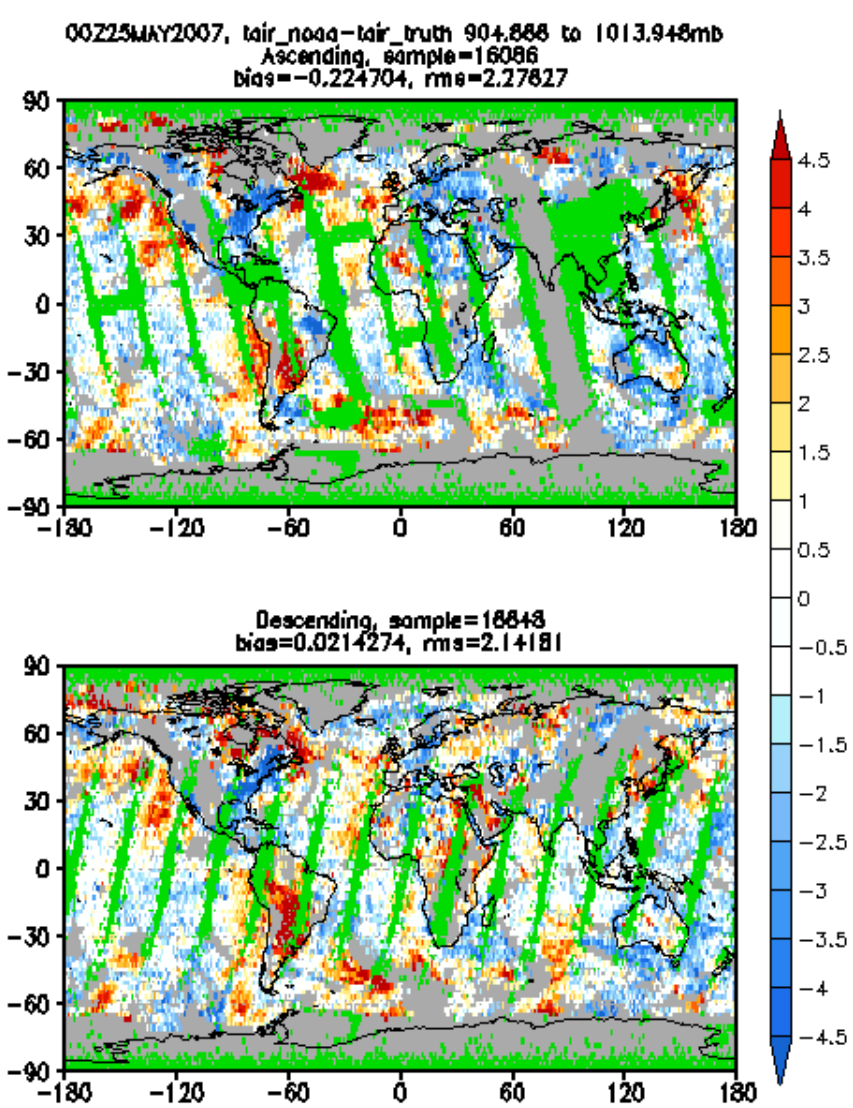


Predictors: AMSU + IASI

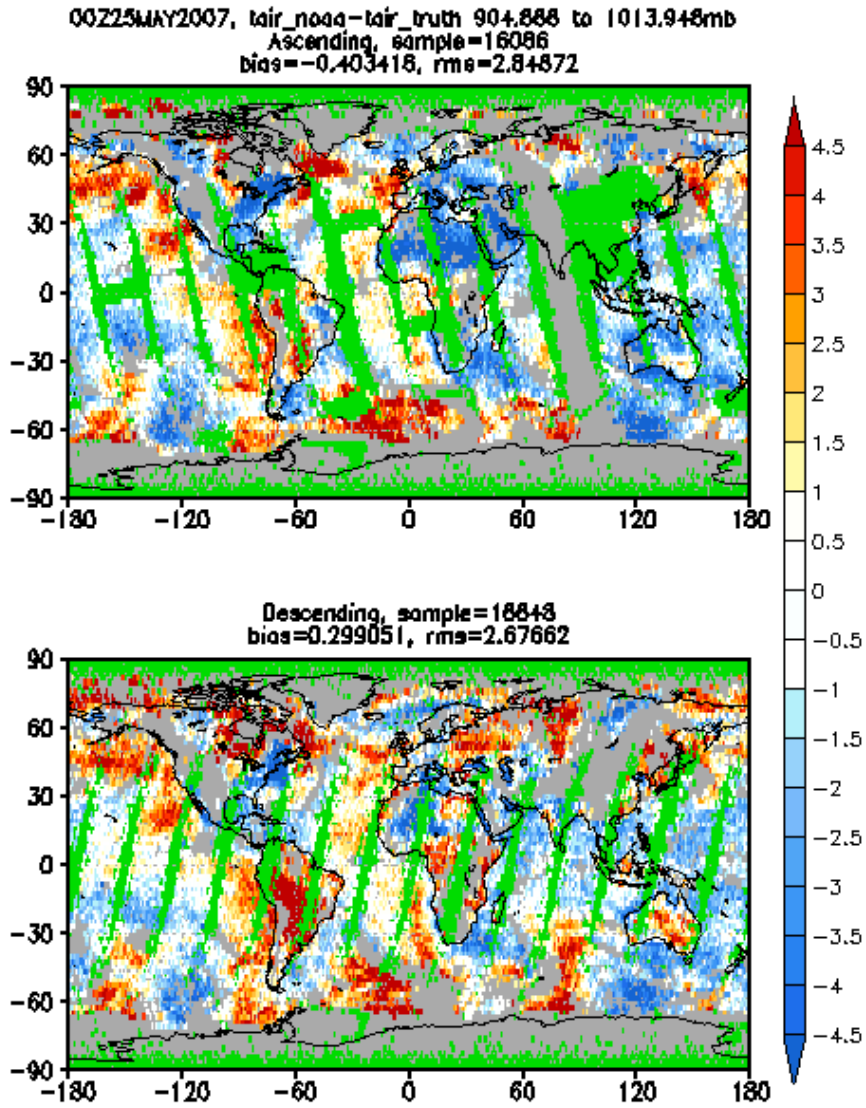


Predictors: AMSU

Regression-ECMWF: Temperature at 904.866 — 1013.948 hPa



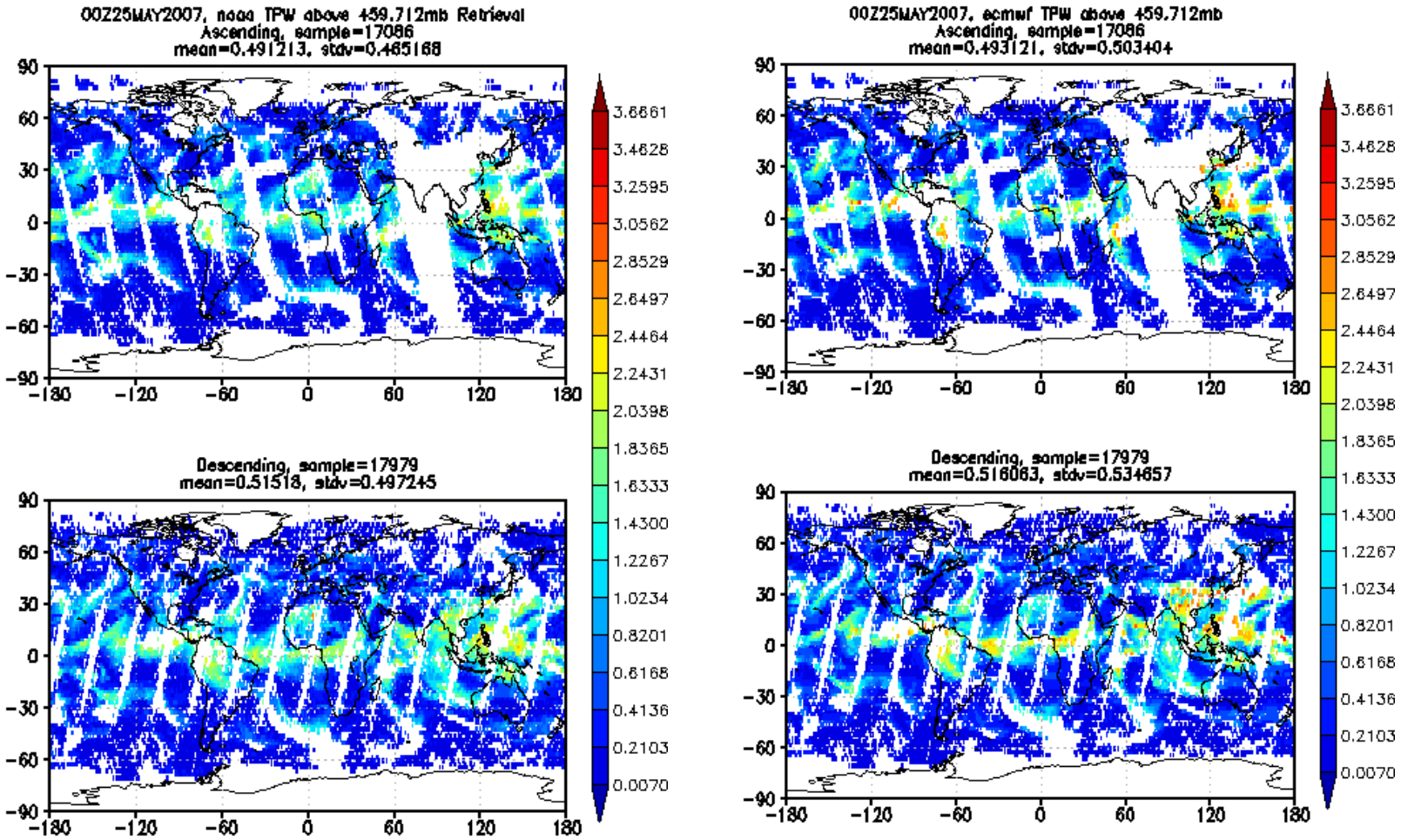
Predictors: AMSU + IASI



Predictors: AMSU

IASI Regression Retrieval—Water Vapor

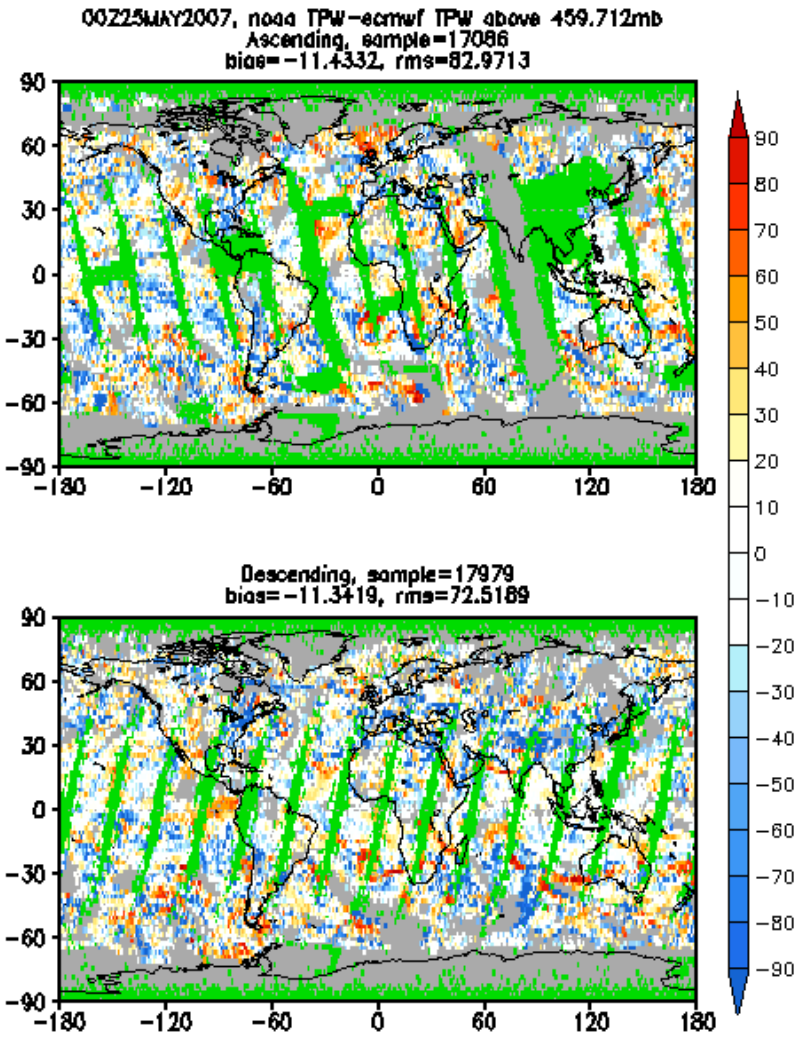
Total Perceptible Water (mm) above 459.712 hPa

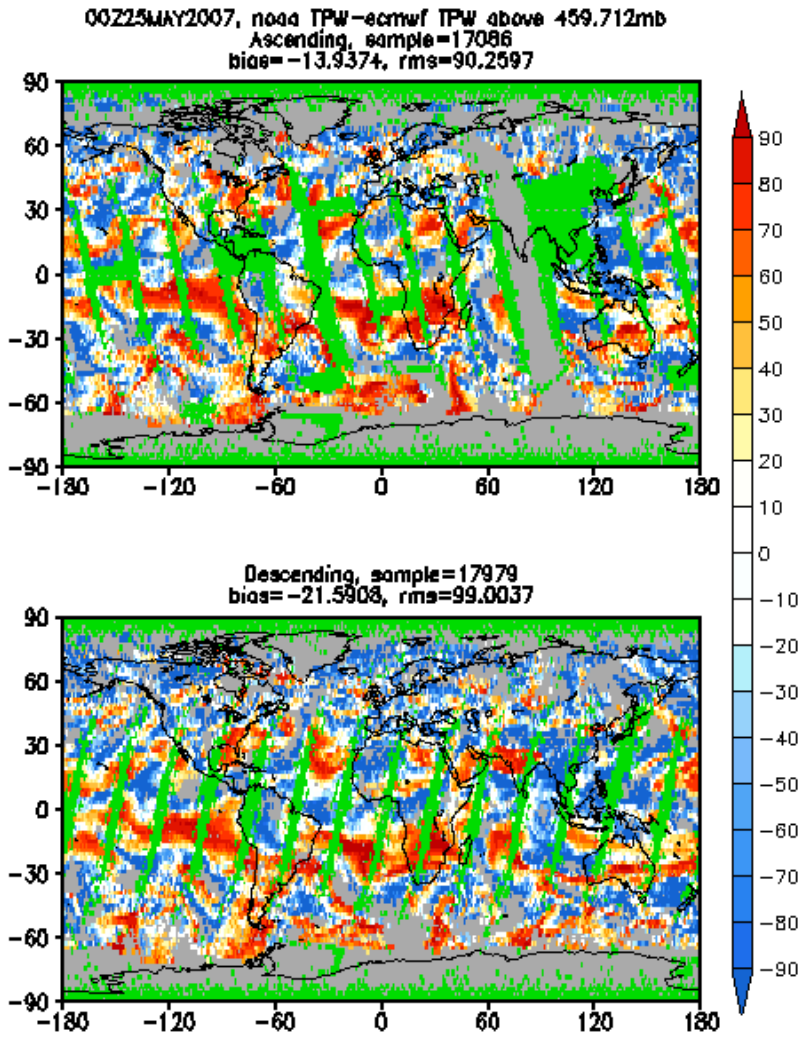


NOAA regression Retrieval
Predictors: AMSU + IASI

ECMWF

Regression – ECMWF TPW Above 459.712 hPa

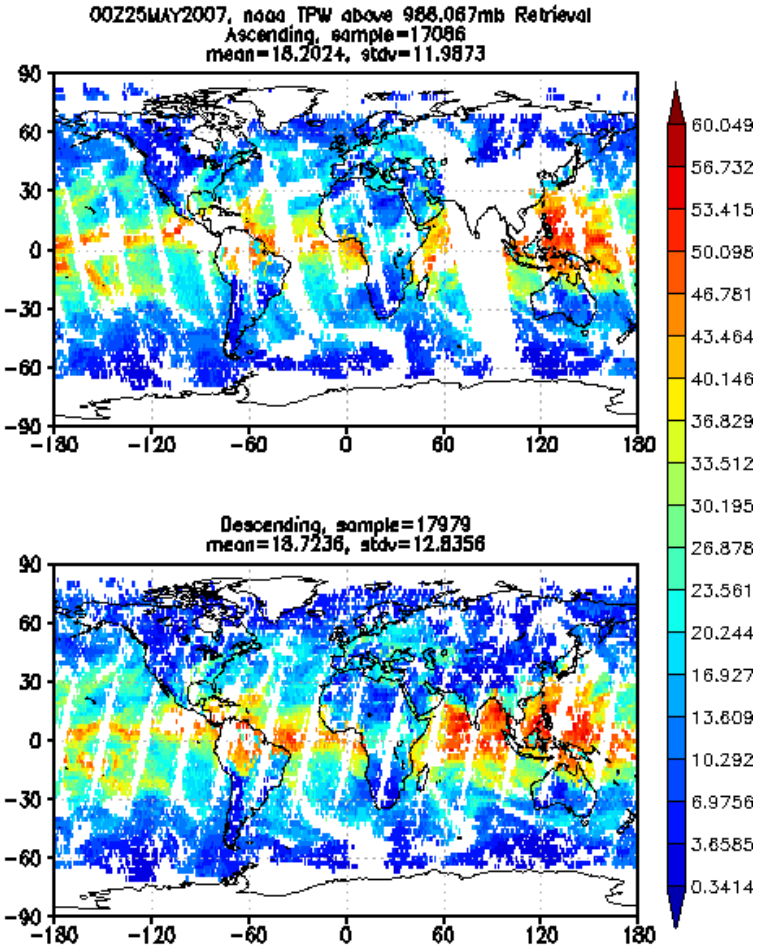




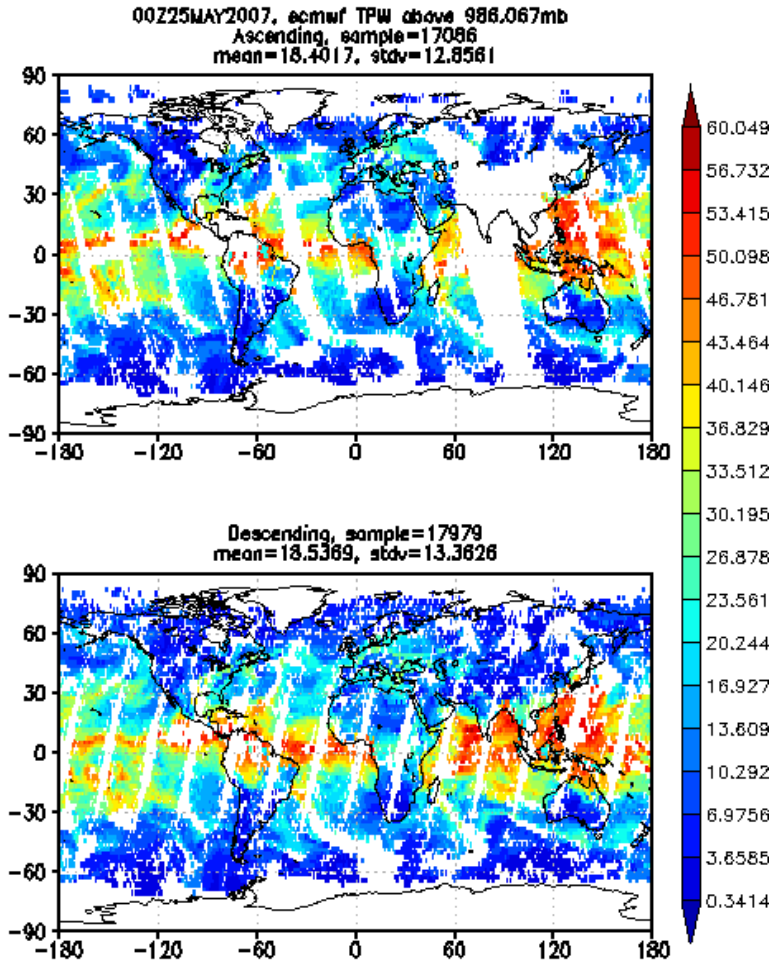
Predictors: AMSU + IASI

Predictors: AMSU (no MHS)

Total Perceptible Water (mm) above 986.067 hPa

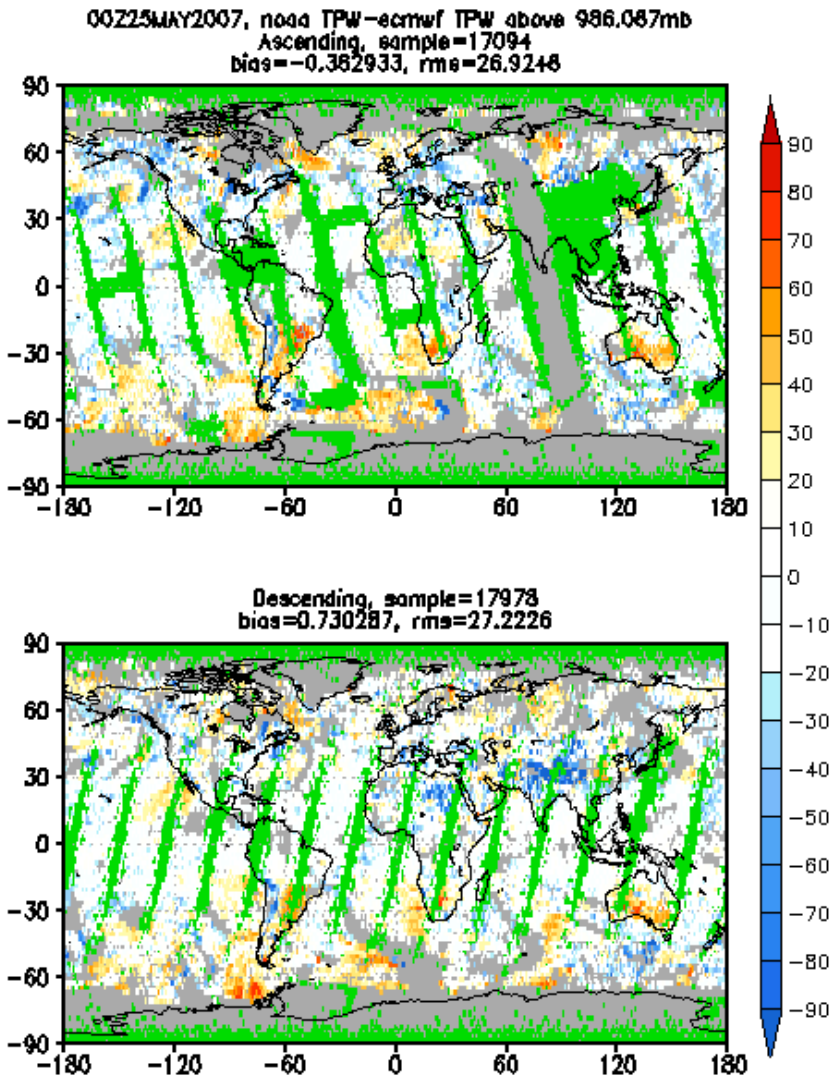


NOAA regression Retrieval
Predictors: AMSU + IASI

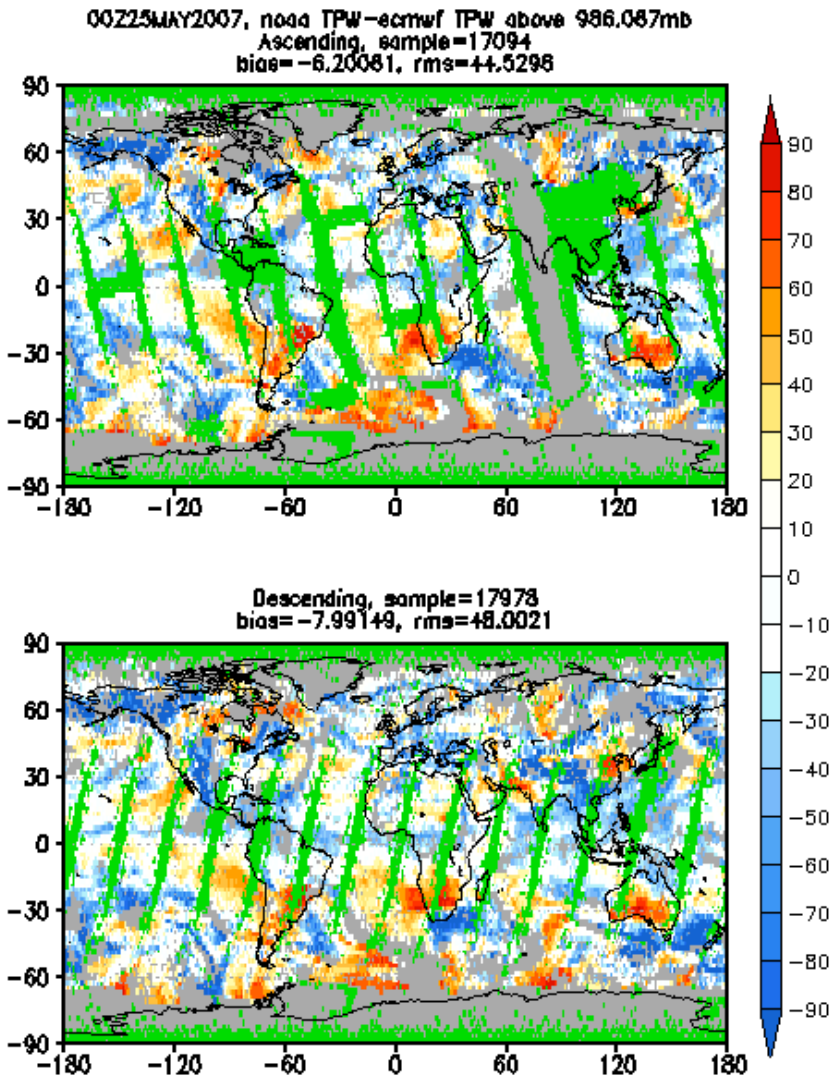


ECMWF

Regression – ECMWF TPW Above 986.067 hPa



Predictors: AMSU + IASI

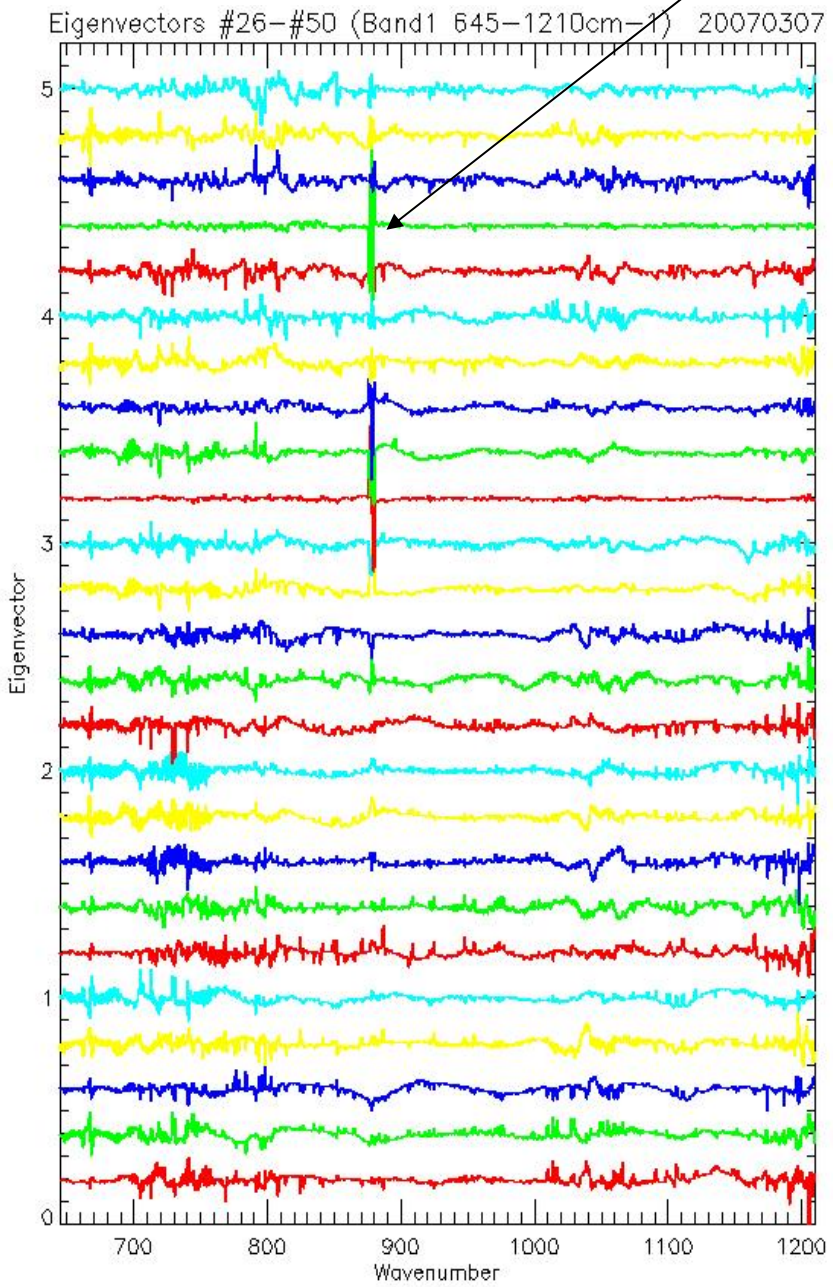


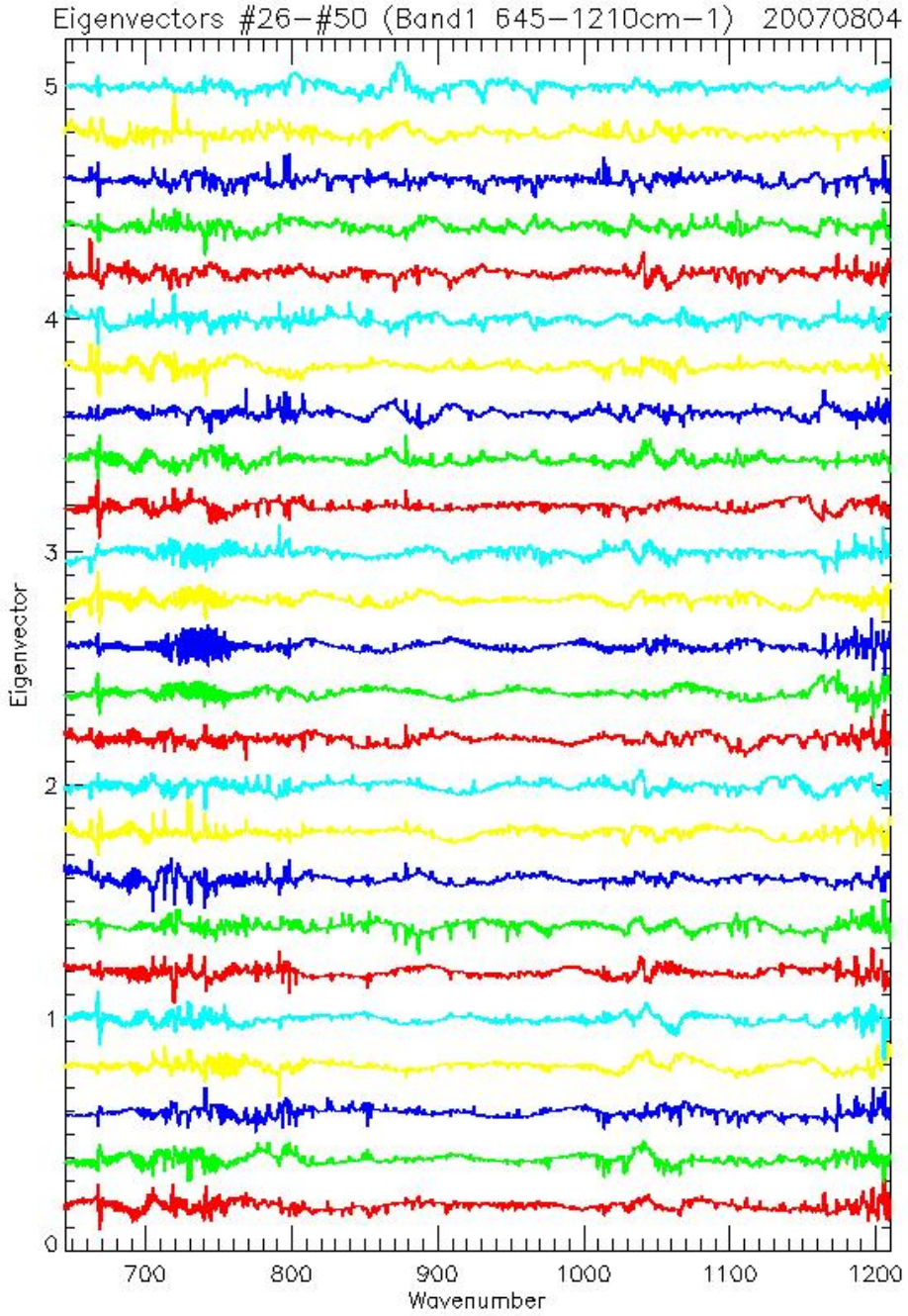
Predictors: AMSU (no MHS)

Summary

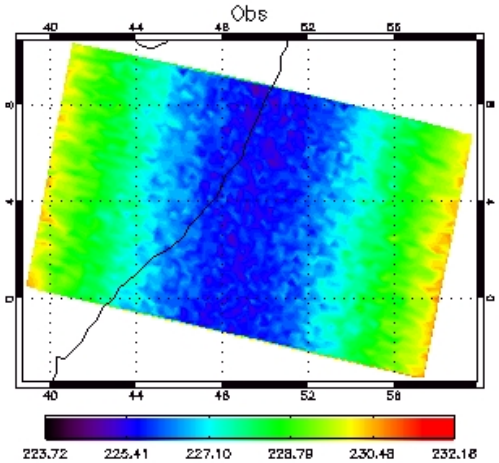
- PCA is working quite well!!
- Static global eigenvectors appear to be representing all situations.
- PCA temperature retrievals performing quite well in partly cloudy/clear conditions
- Plan to test in a data assimilation experiment.
- If anyone else is interested please email.
  - mitch.goldberg@noaa.gov
  - Generating the PCA retrievals from the NESDIS Operational PC BUFR files is very straightforward.

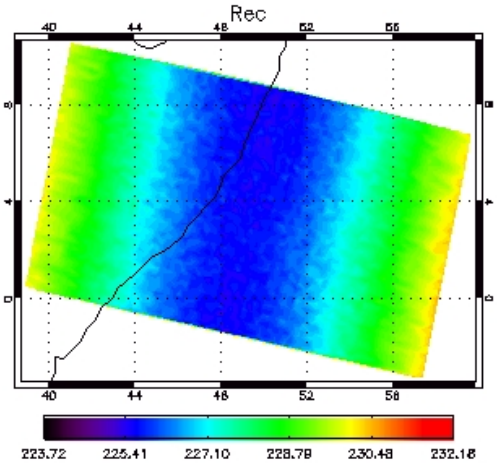
Odd features in older data (March 2007)

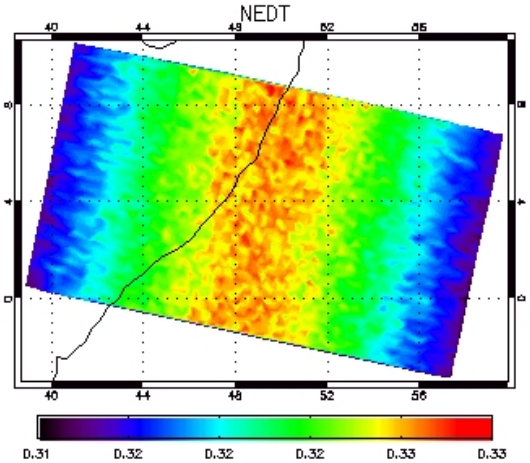


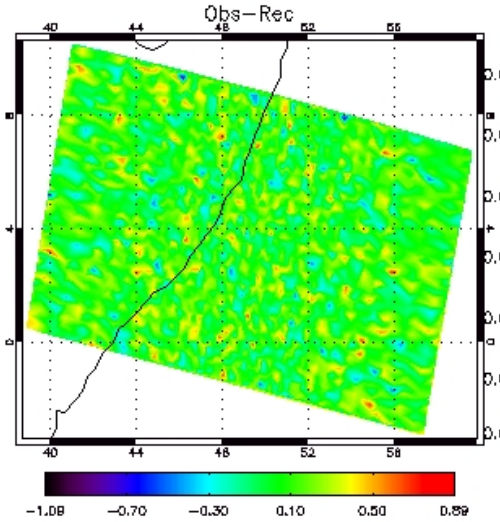


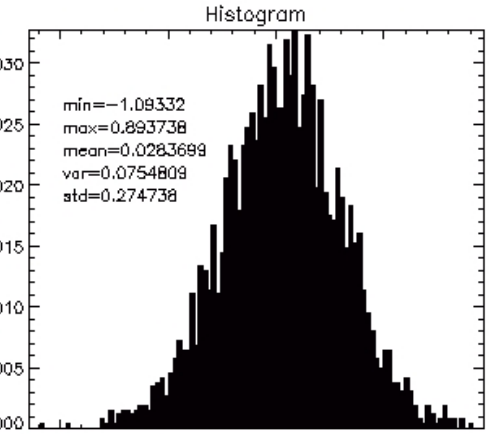
- Additional Slides showing other selected channels for pages 22-26

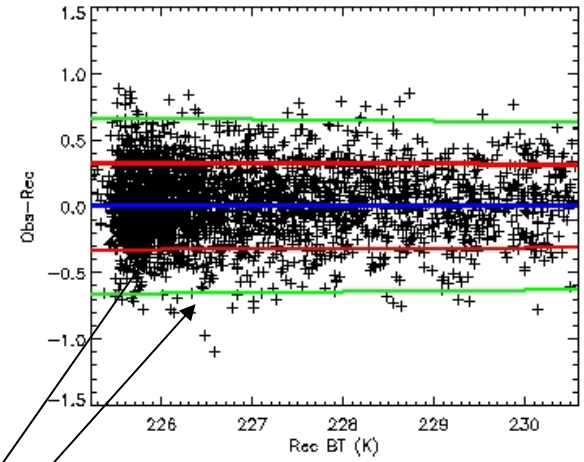






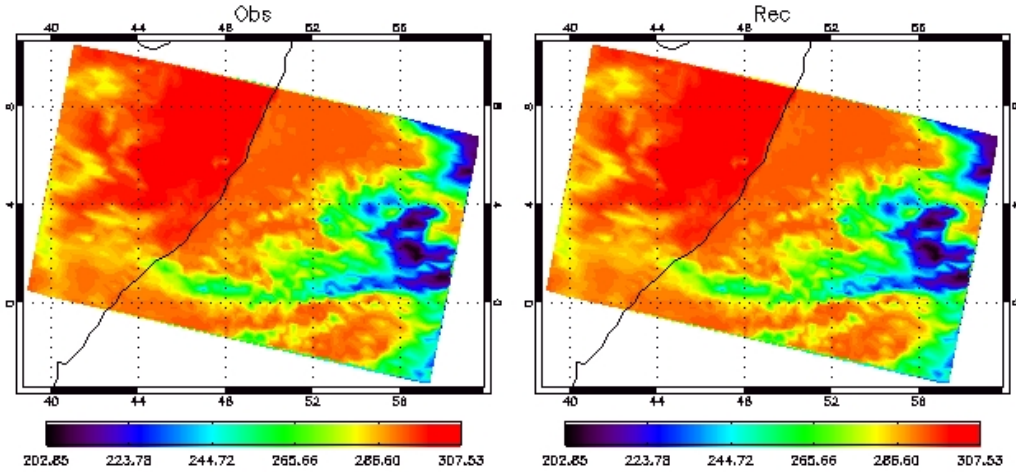


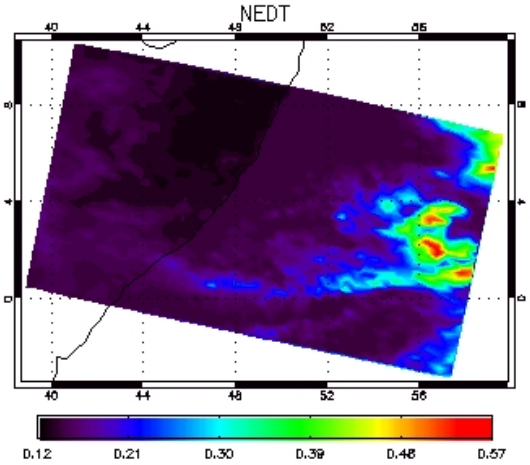


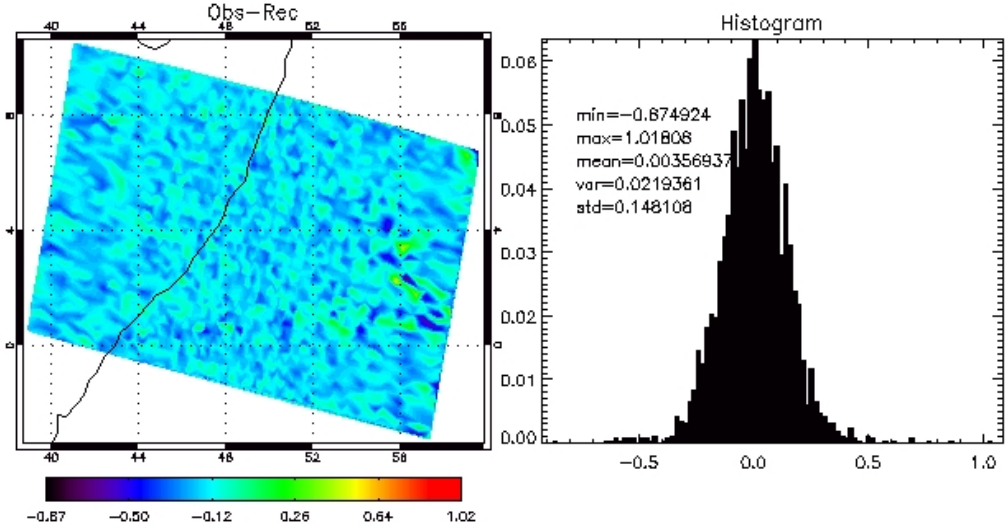


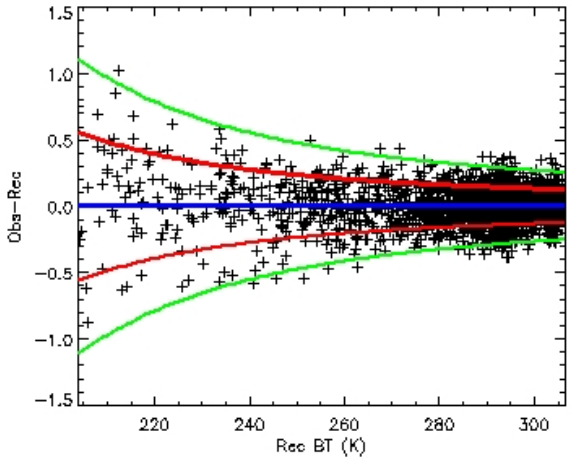
IASI [670.000cm−1] Granule−125 20070910

IASI [670.000cm−1] Granule−125 20070910



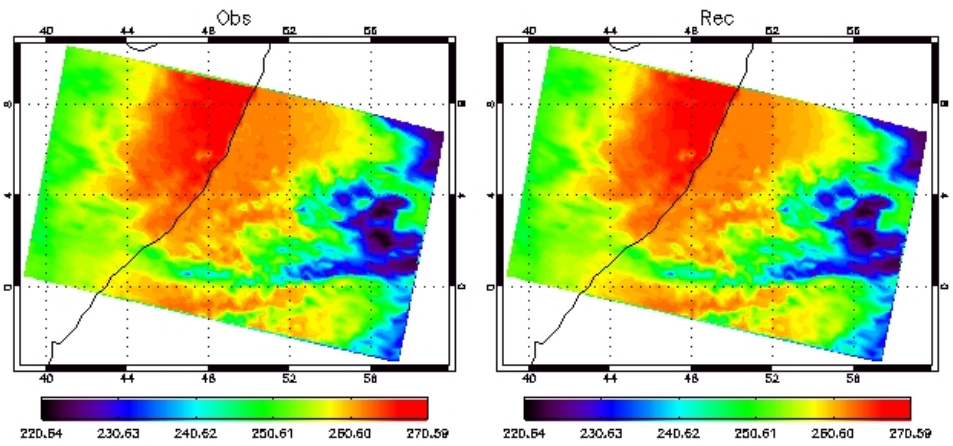


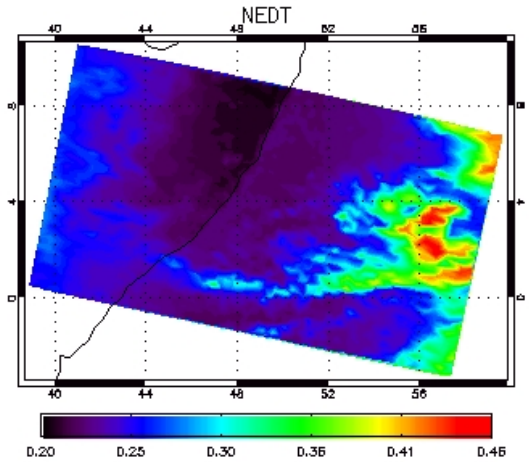


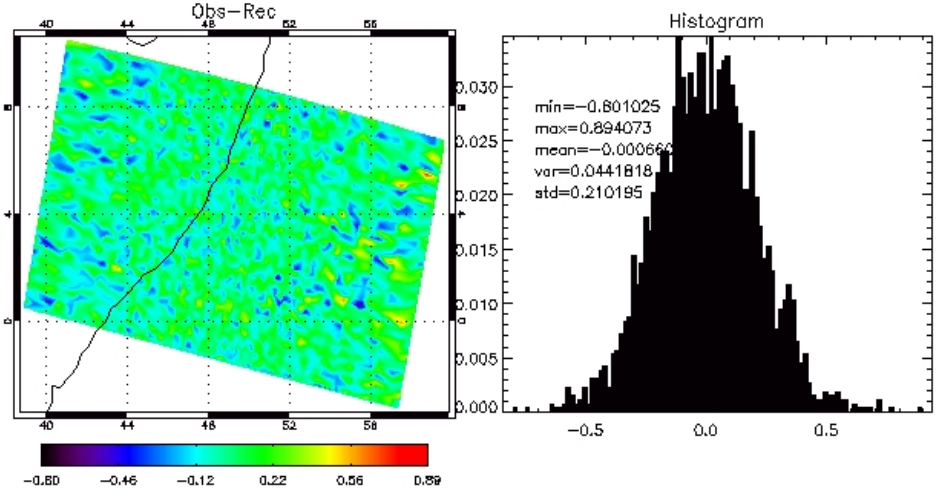


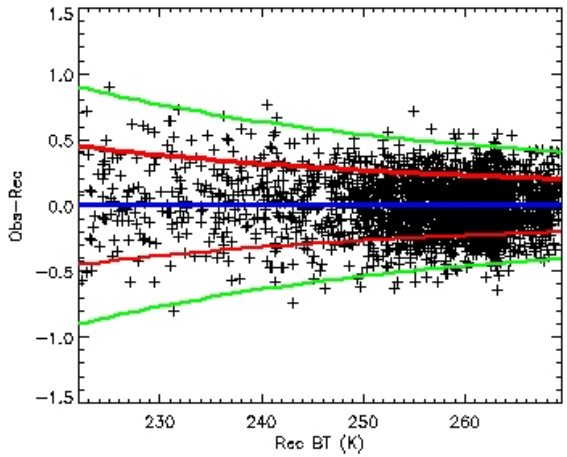
IASI [965.000cm−1] Granule−125 20070910

IASI [965.000cm−1] Granule−125 20070910



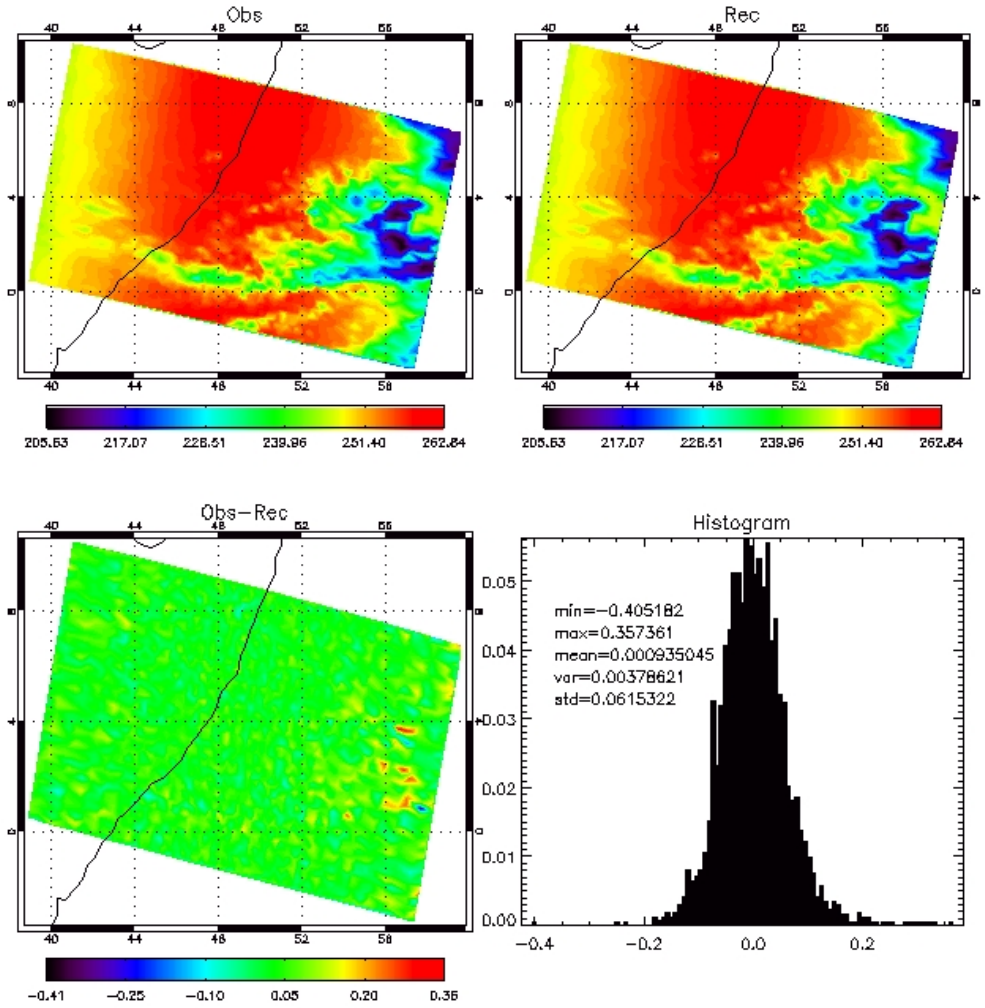






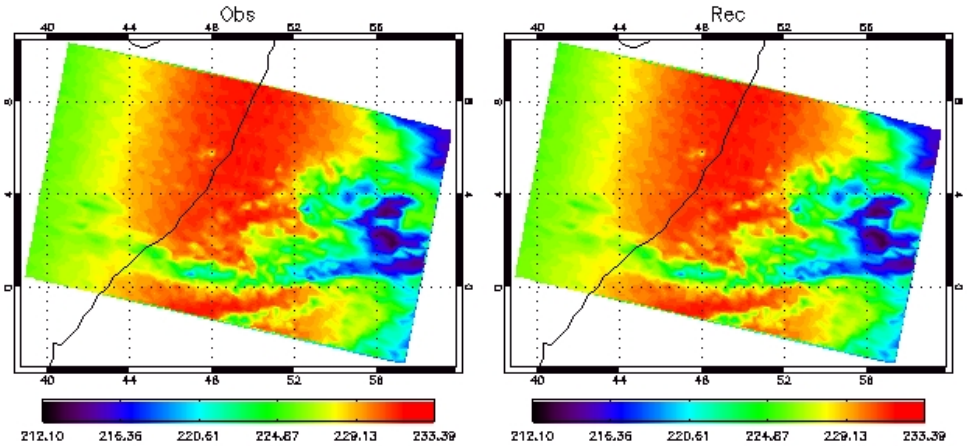
IASI [1041.50cm−1] Granule−125 20070910

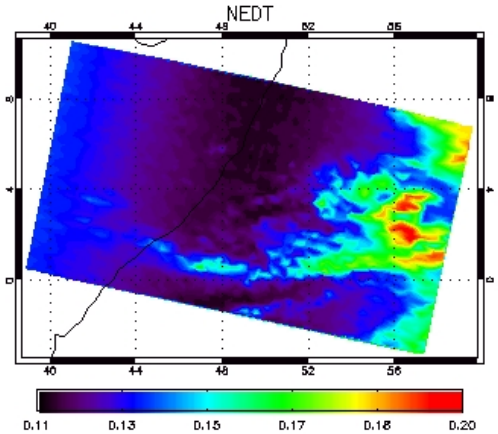
IASI [1041.50cm−1] Granule−125 20070910

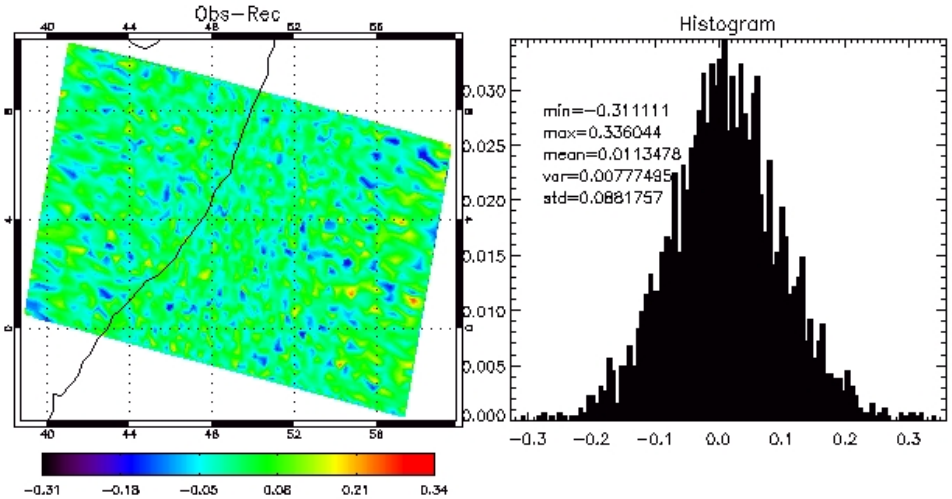


IASI [1298.00cm−1] Granule−125 20070910

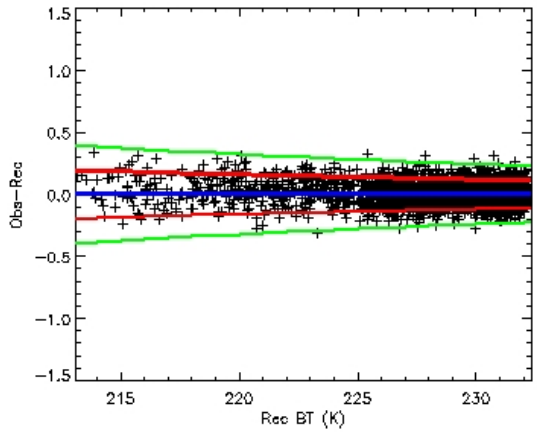
IASI [1298.00cm−1] Granule−125 20070910



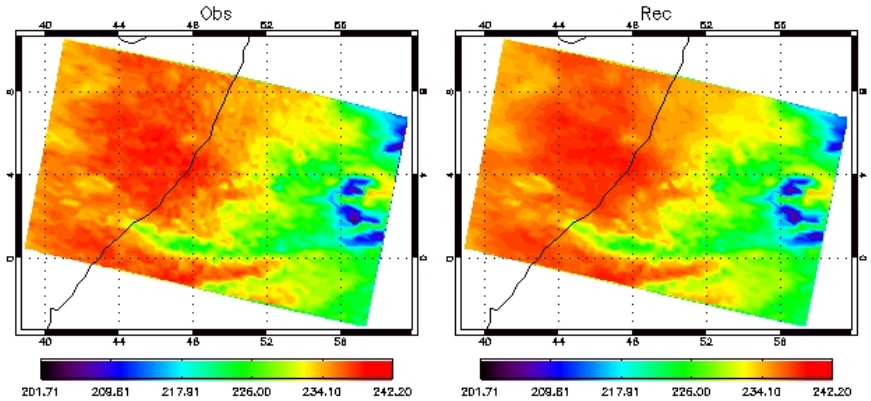


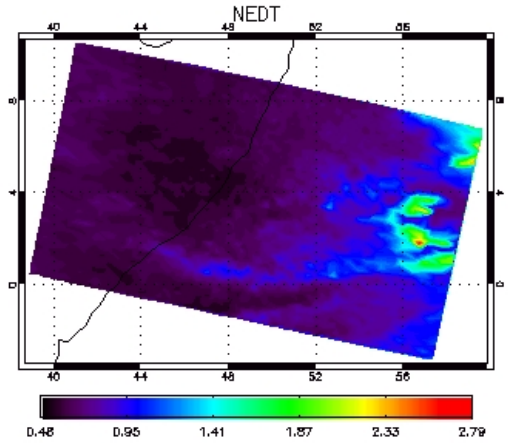


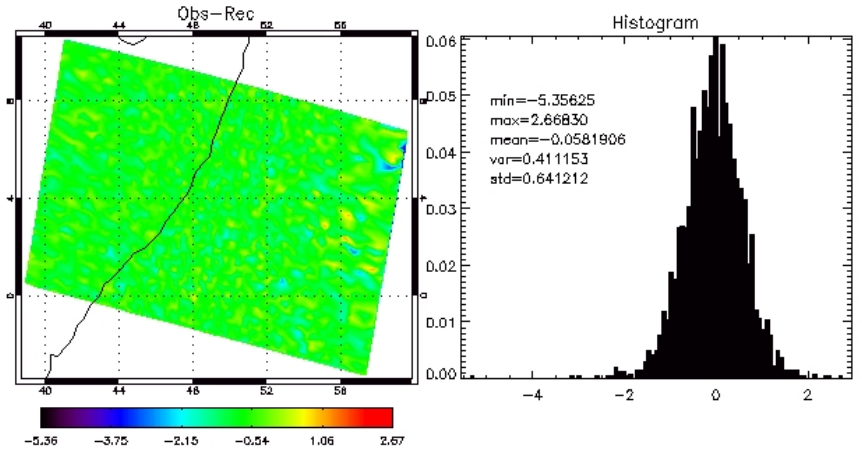
IASI [1306.25cm−1] Granule−125 20070910

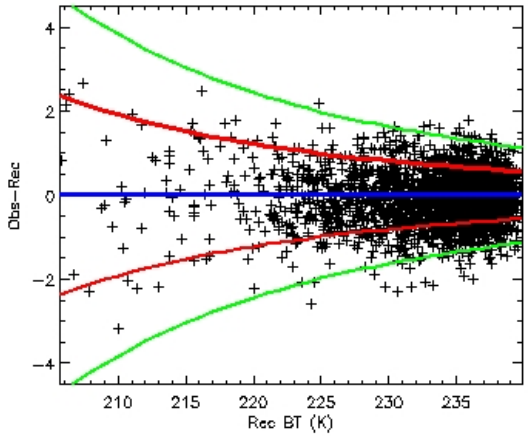


IASI [1306.25cm−1] Granule−125 20070910



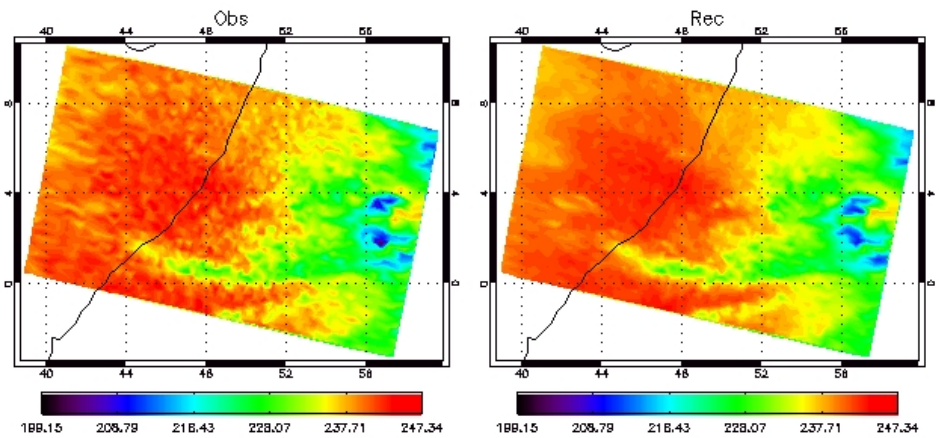


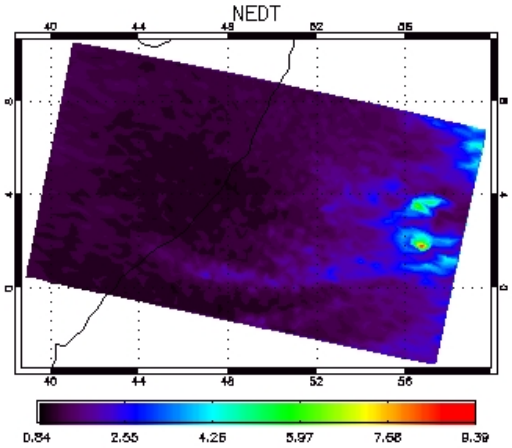


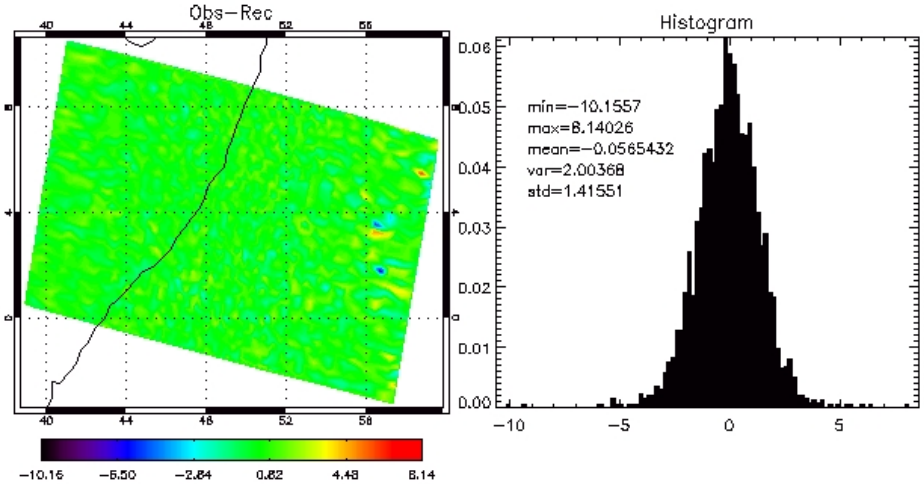


IASI [1739.00cm−1] Granule−125 20070910

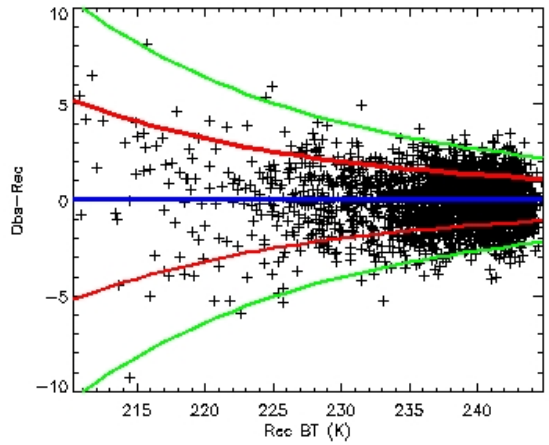
IASI [1739.00cm−1] Granule−125 20070910



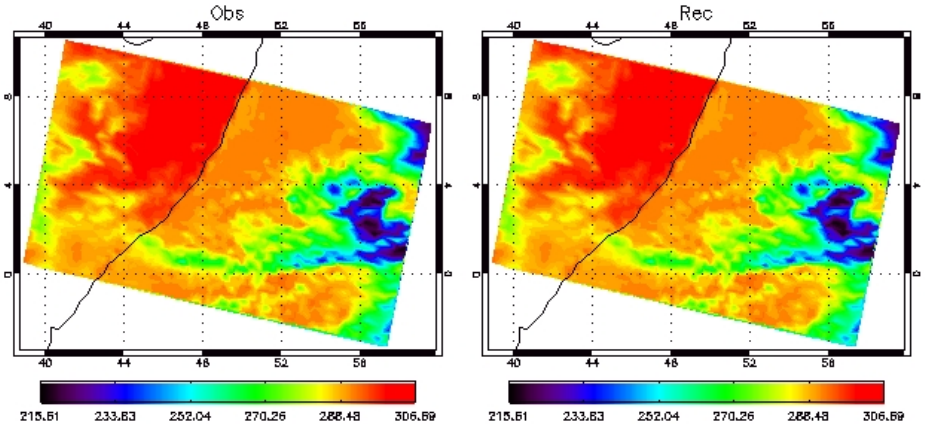


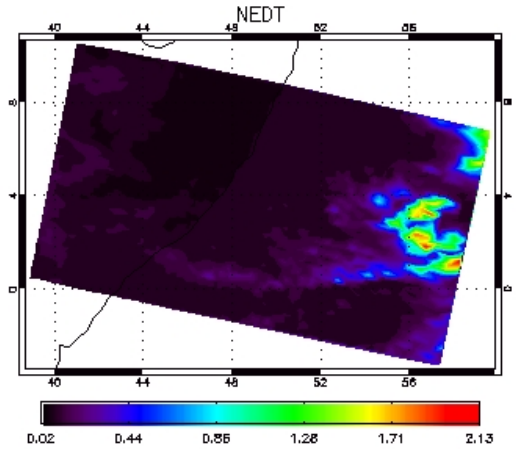


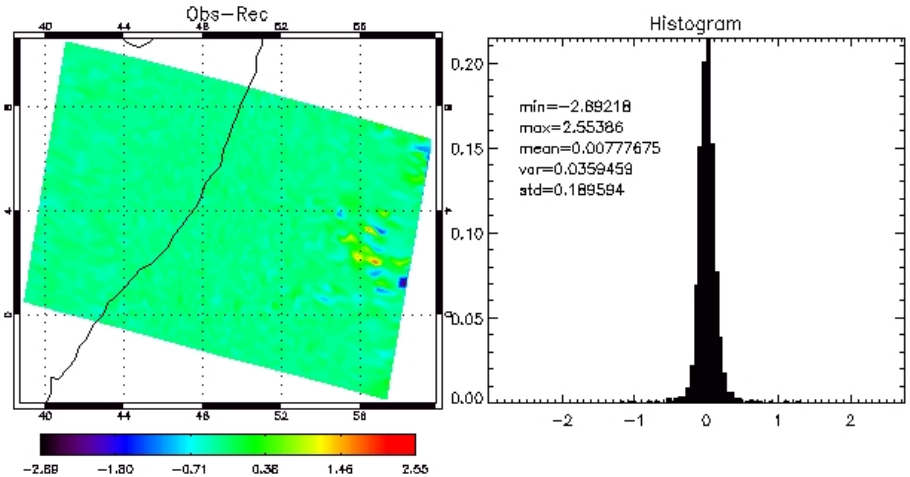
IASI [1923.00cm−1] Granule−125 20070910



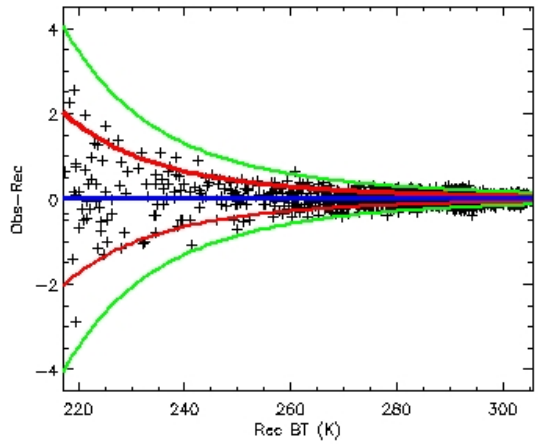
IASI [1923.00cm−1] Granule−125 20070910



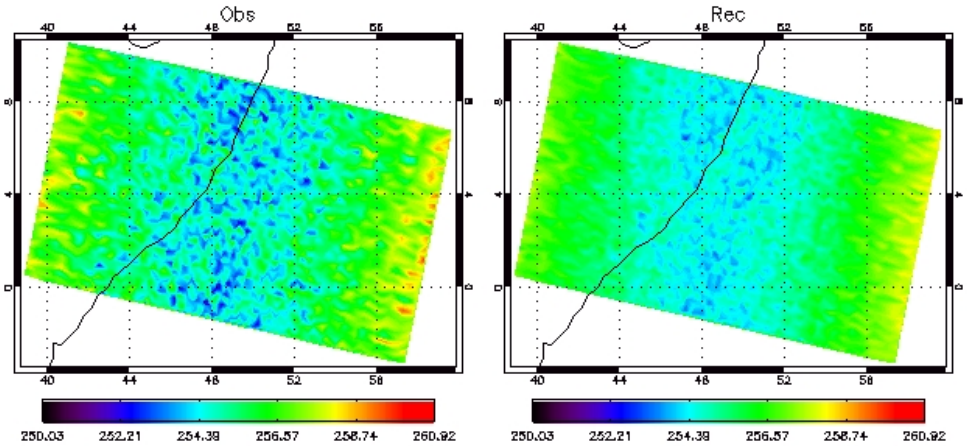


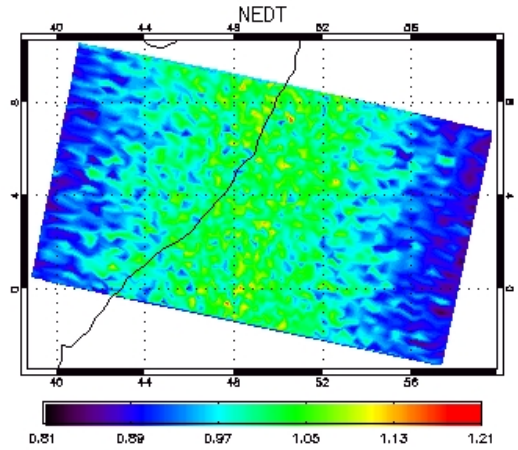


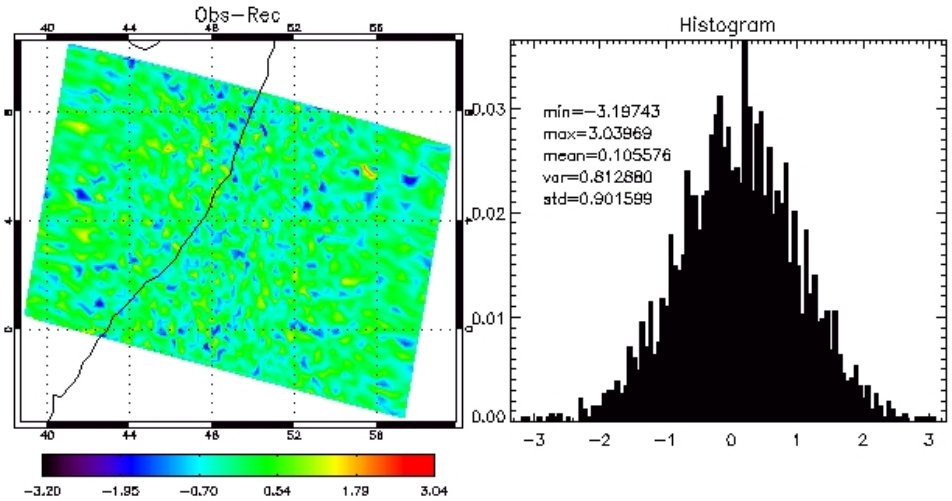
IASI [2105cm−1] Granule−125 20070910



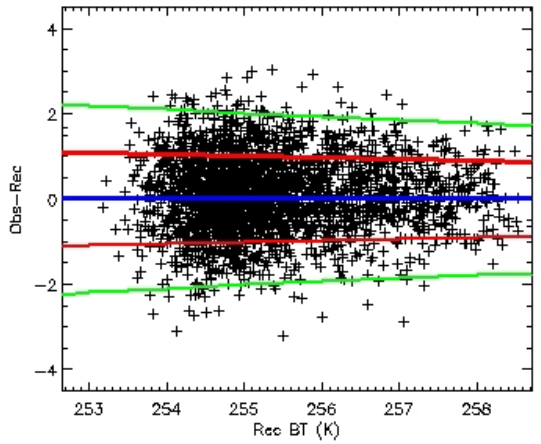
IASI [2105cm−1] Granule−125 20070910



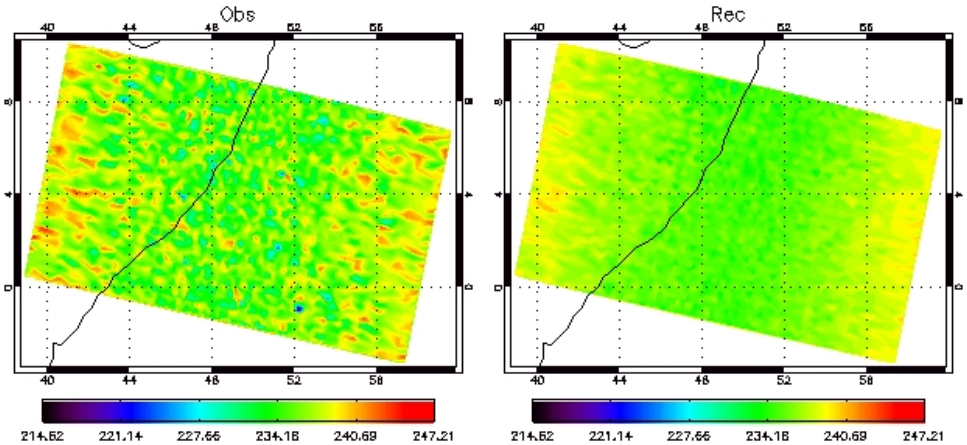


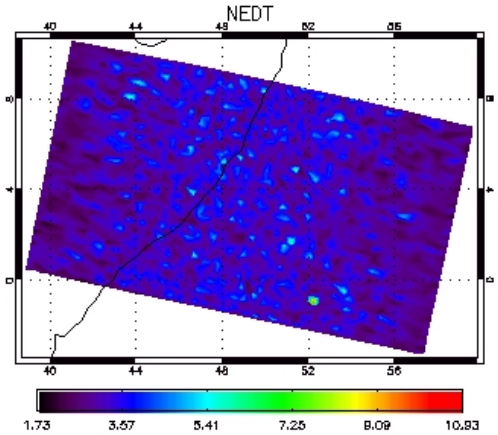


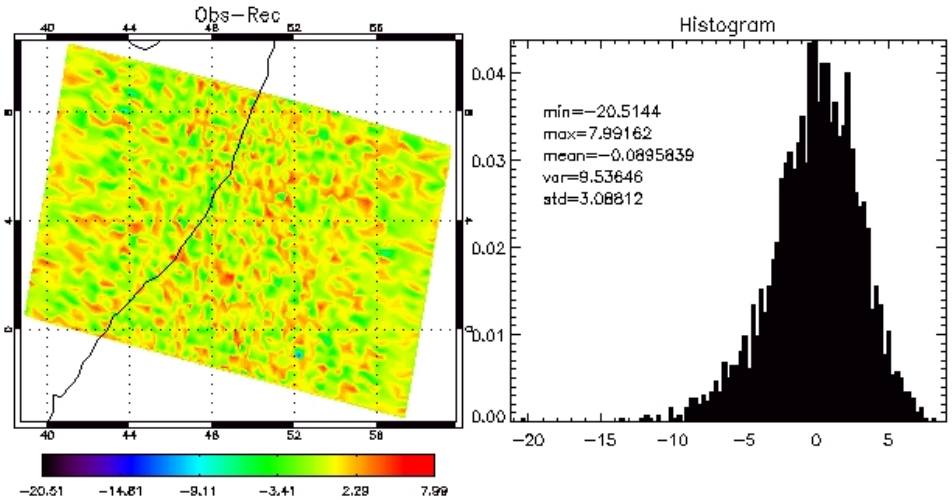
IASI [2325cm−1] Granule−125 20070910



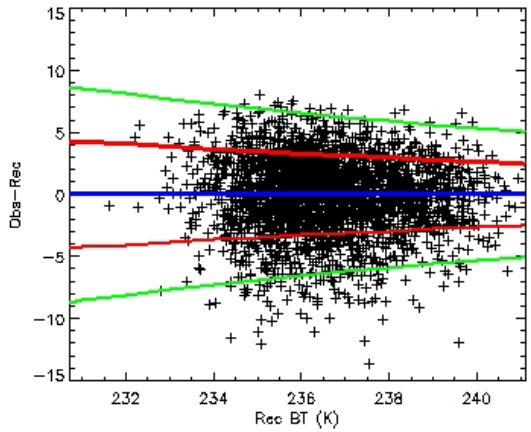
IASI [2325cm−1] Granule−125 20070910



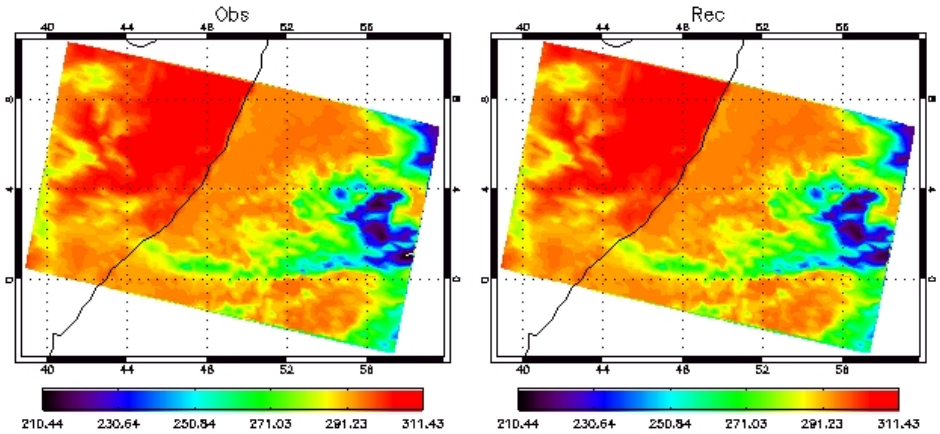


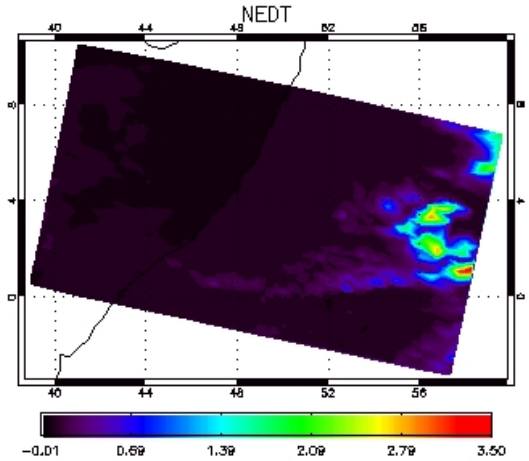


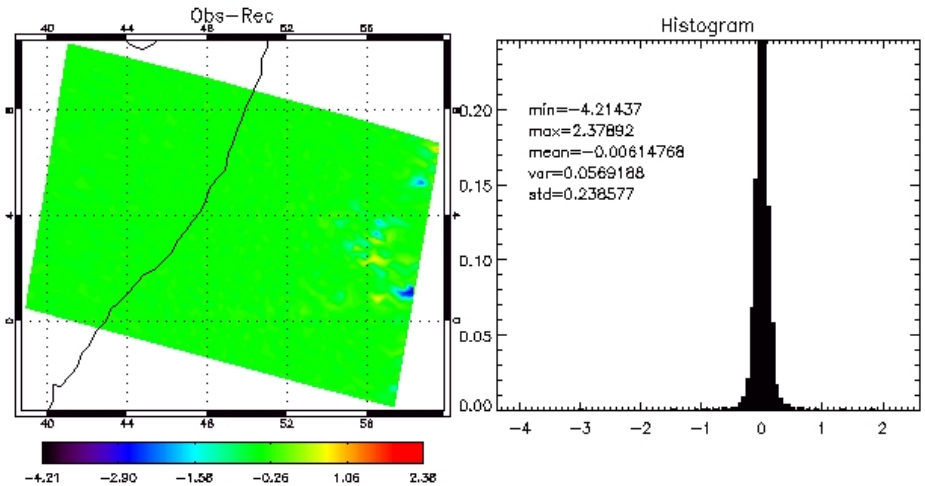
IASI [2375cm−1] Granule−125 20070910

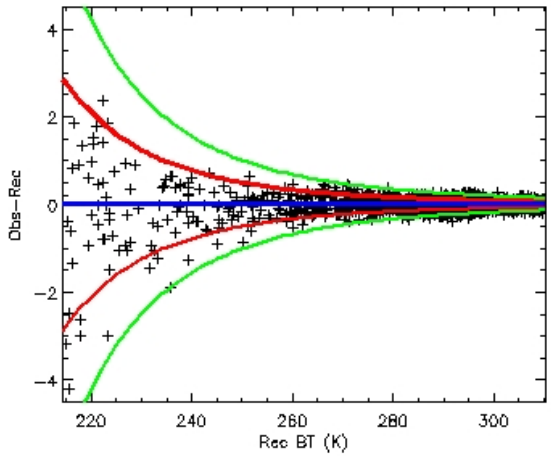


IASI [2375cm−1] Granule−125 20070910







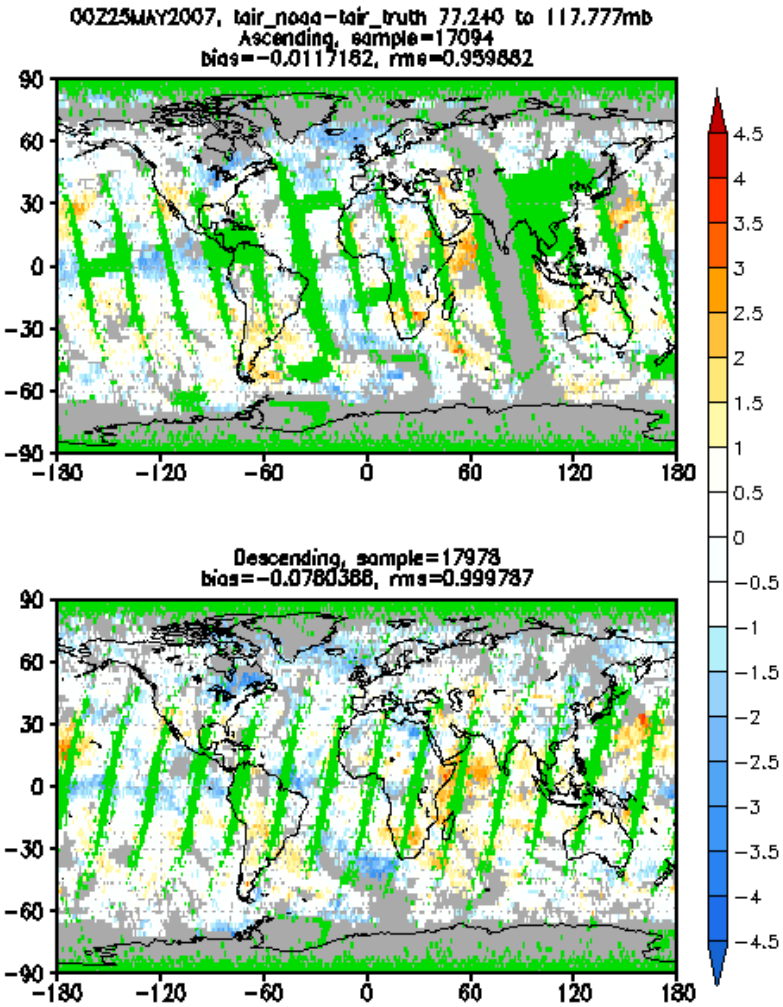


IASI [2141cm−1] Granule−125 20070910

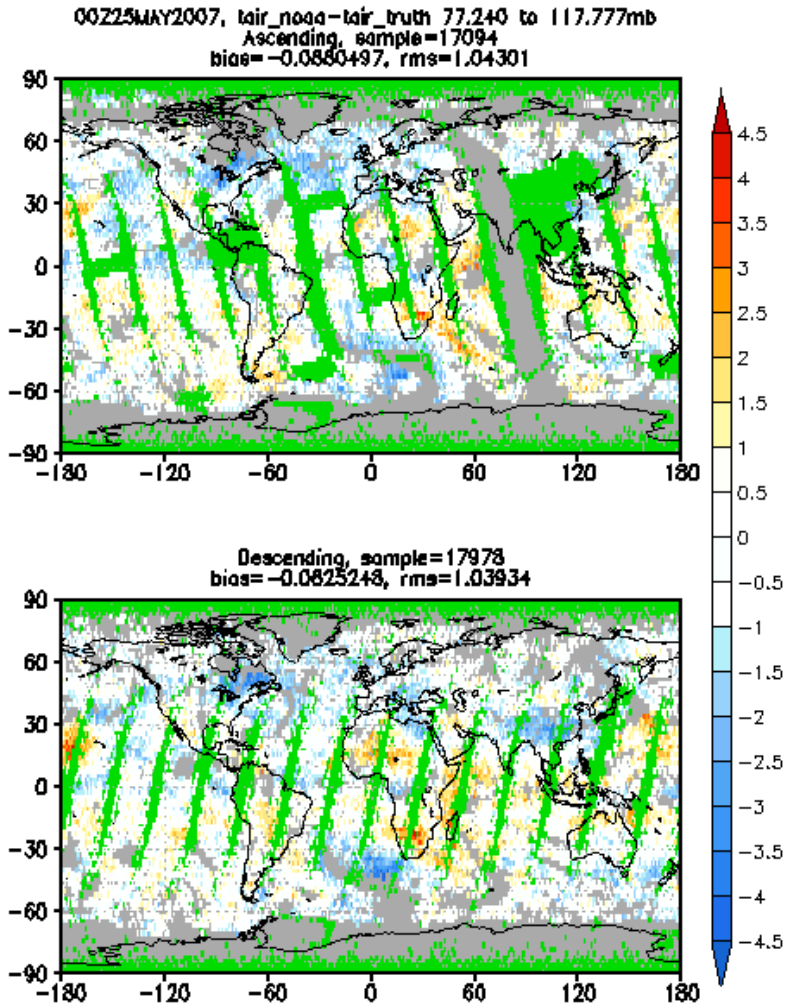
IASI [2141cm−1] Granule−125 20070910

- Additional slides comparing the regression retrieval for temperature (from page 40)

Regression-ECMWF: Temperature at 77.24 — 117.777 hPa

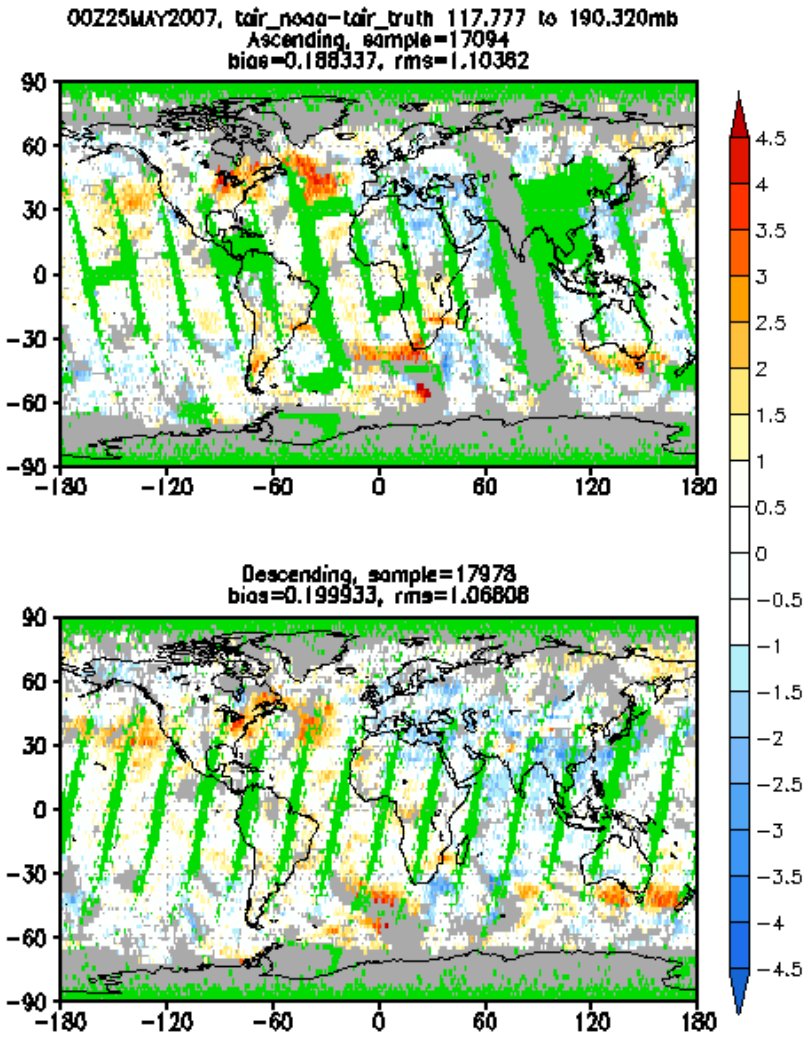


Predictors: AMSU + IASI

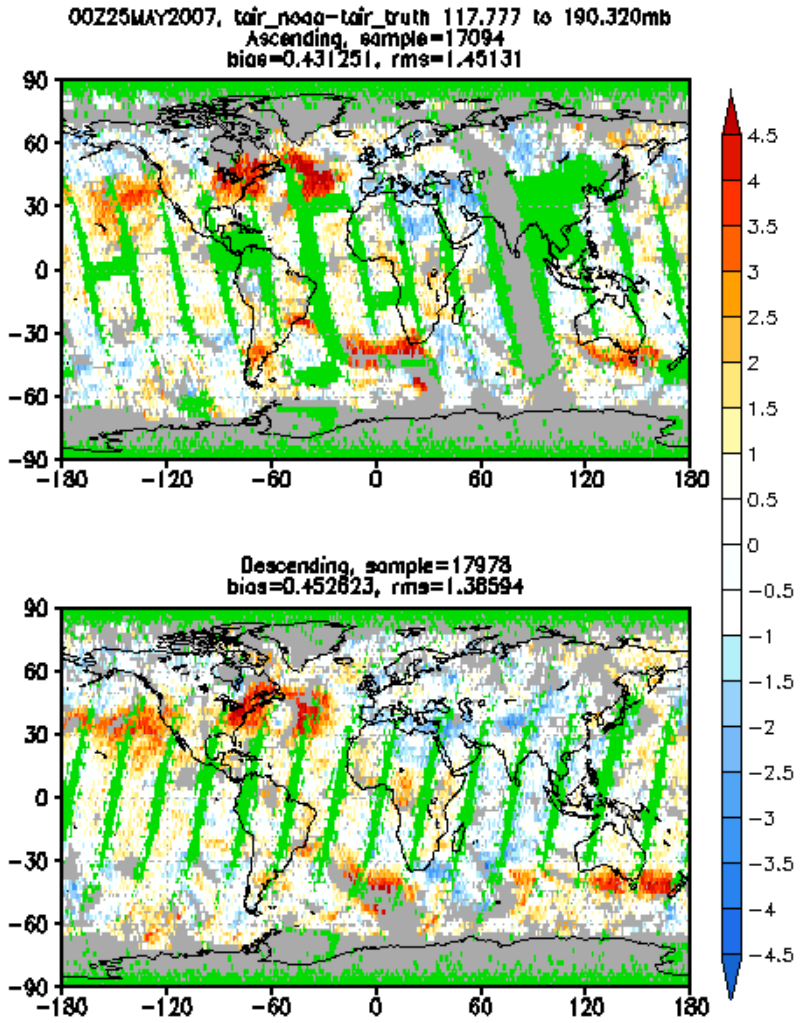


Predictors: AMSU

Regression-ECMWF: Temperature at 117.777 — 190.32 hPa

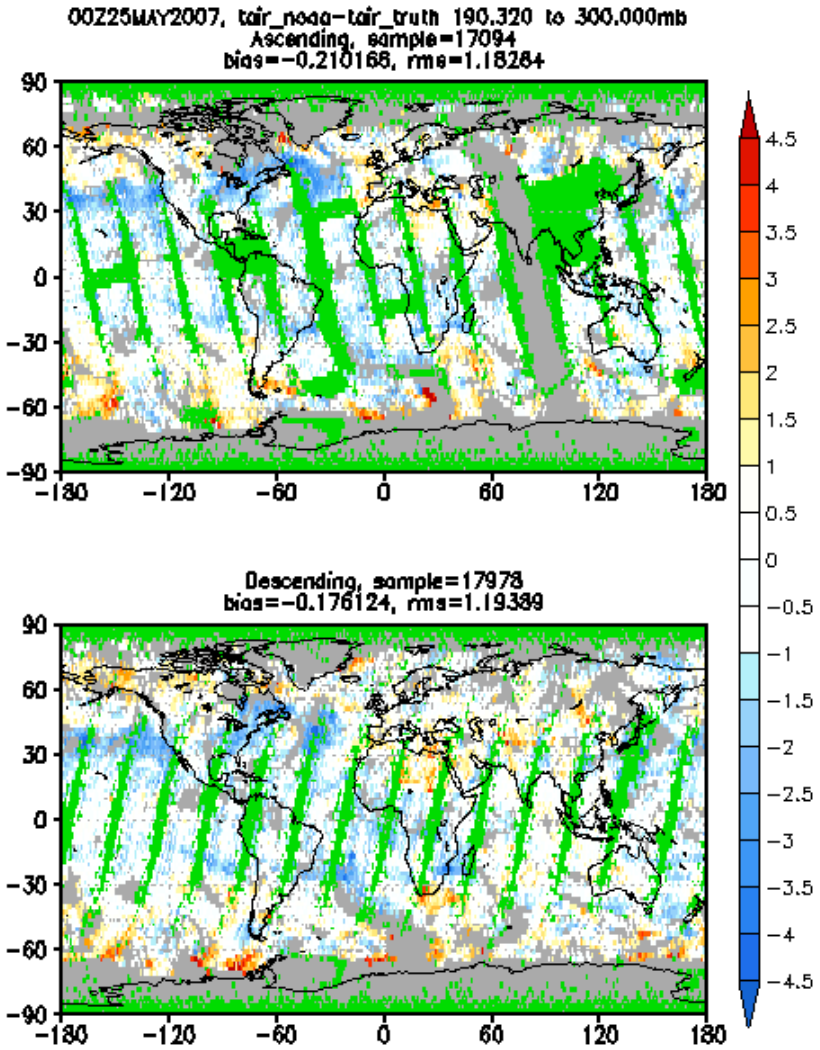


Predictors: AMSU + IASI

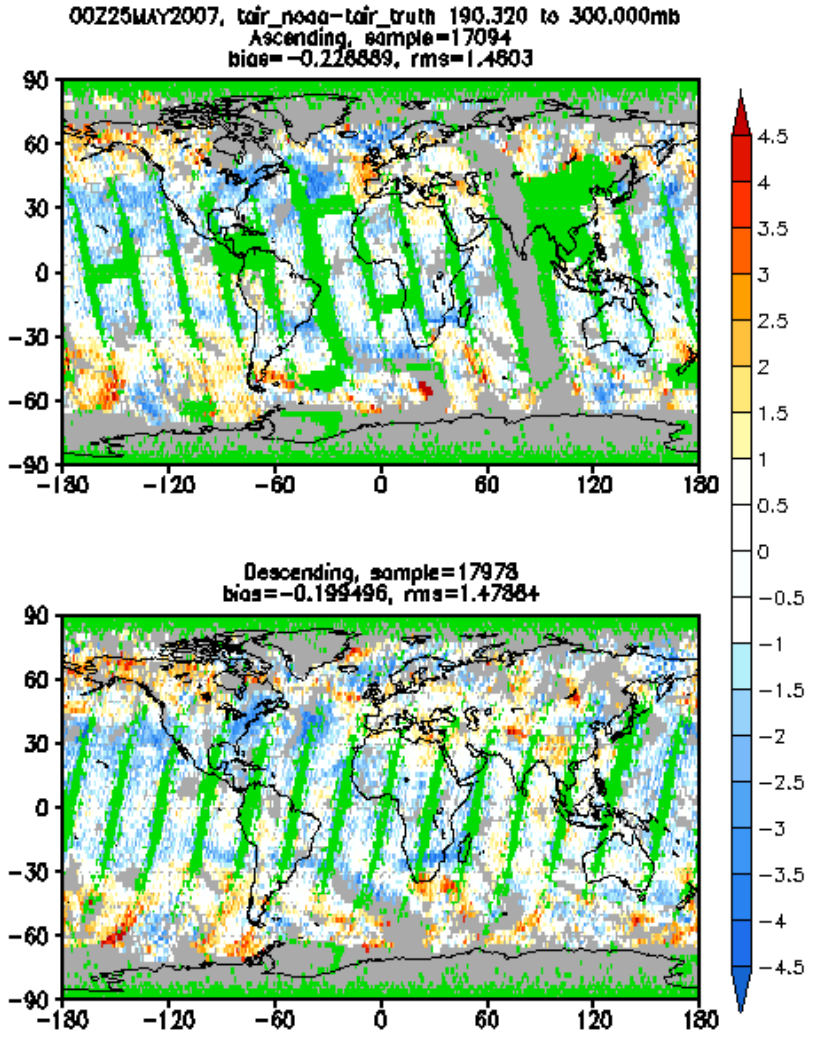


Predictors: AMSU

Regression-ECMWF: Temperature at 190.32 — 300 hPa

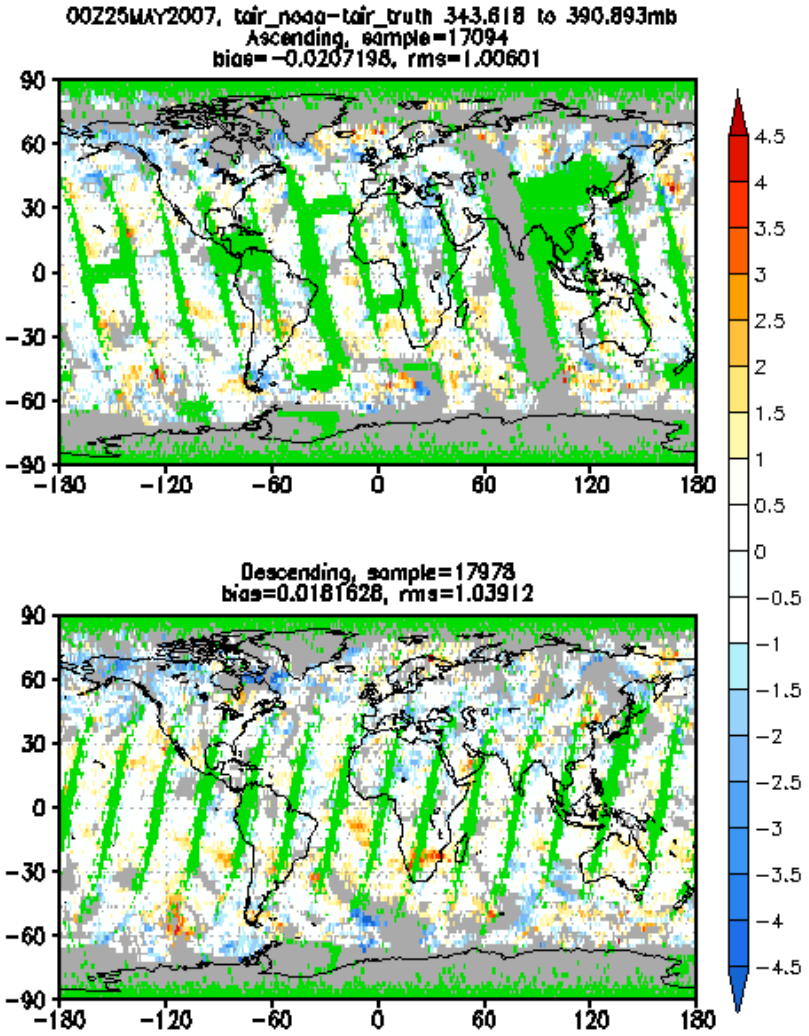


Predictors: AMSU + IASI

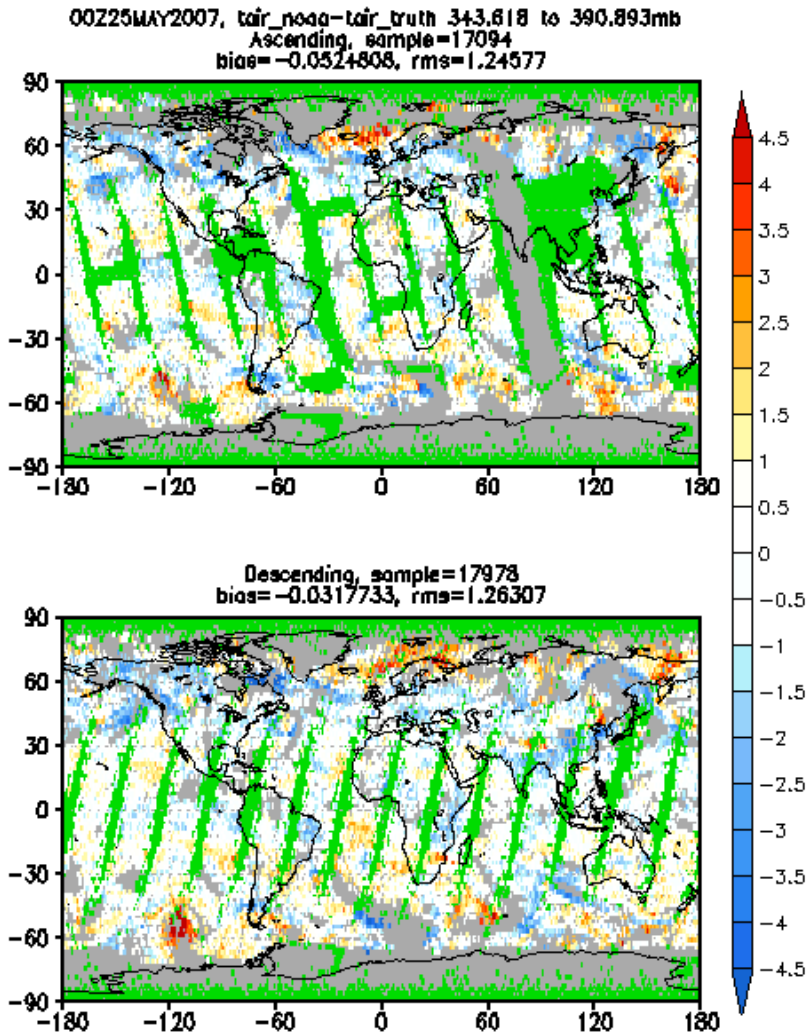


Predictors: AMSU

Regression-ECMWF: Temperature at 343.618 — 390.893 hPa

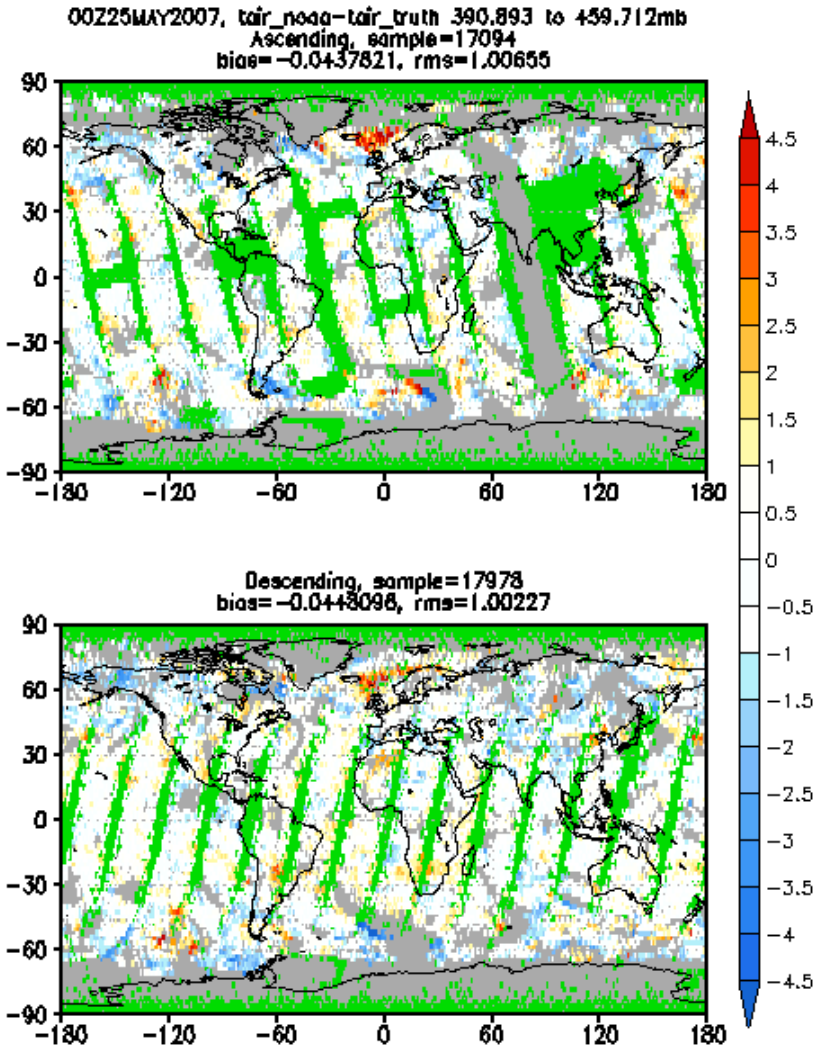


Predictors: AMSU + IASI

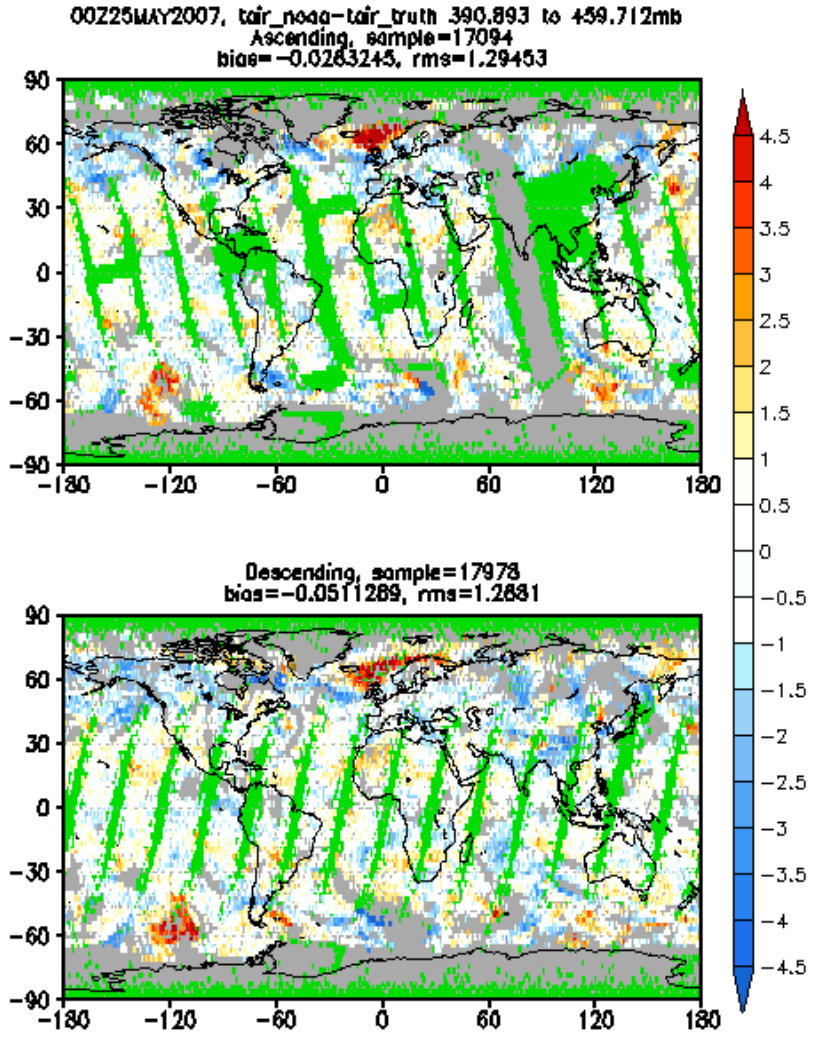


Predictors: AMSU

Regression-ECMWF: Temperature at 390.893 — 459.712 hPa

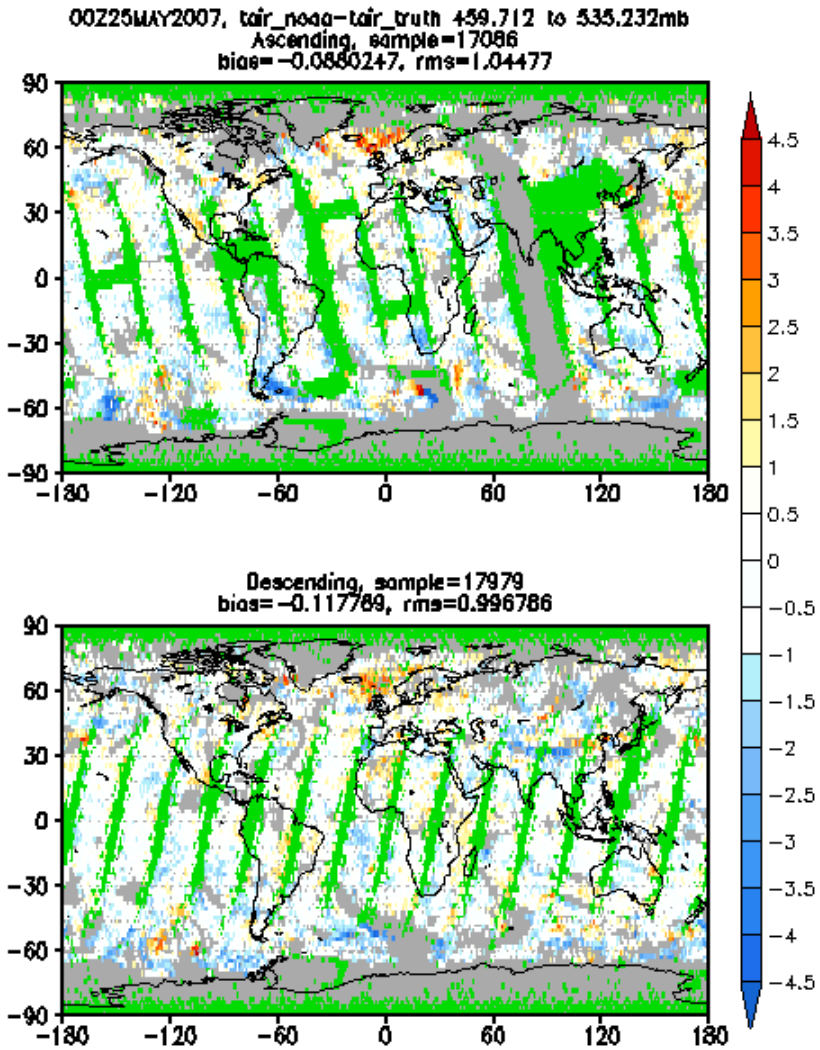


Predictors: AMSU + IASI

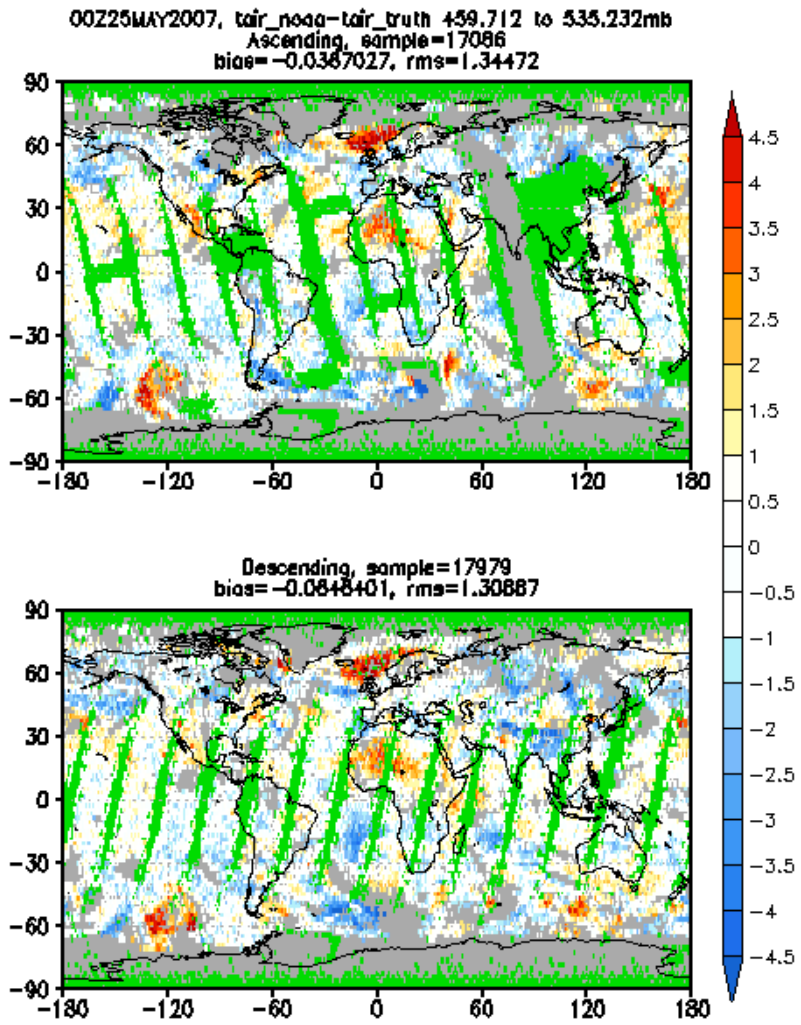


Predictors: AMSU

Regression-ECMWF: Temperature at 459.712 — 535.232 hPa

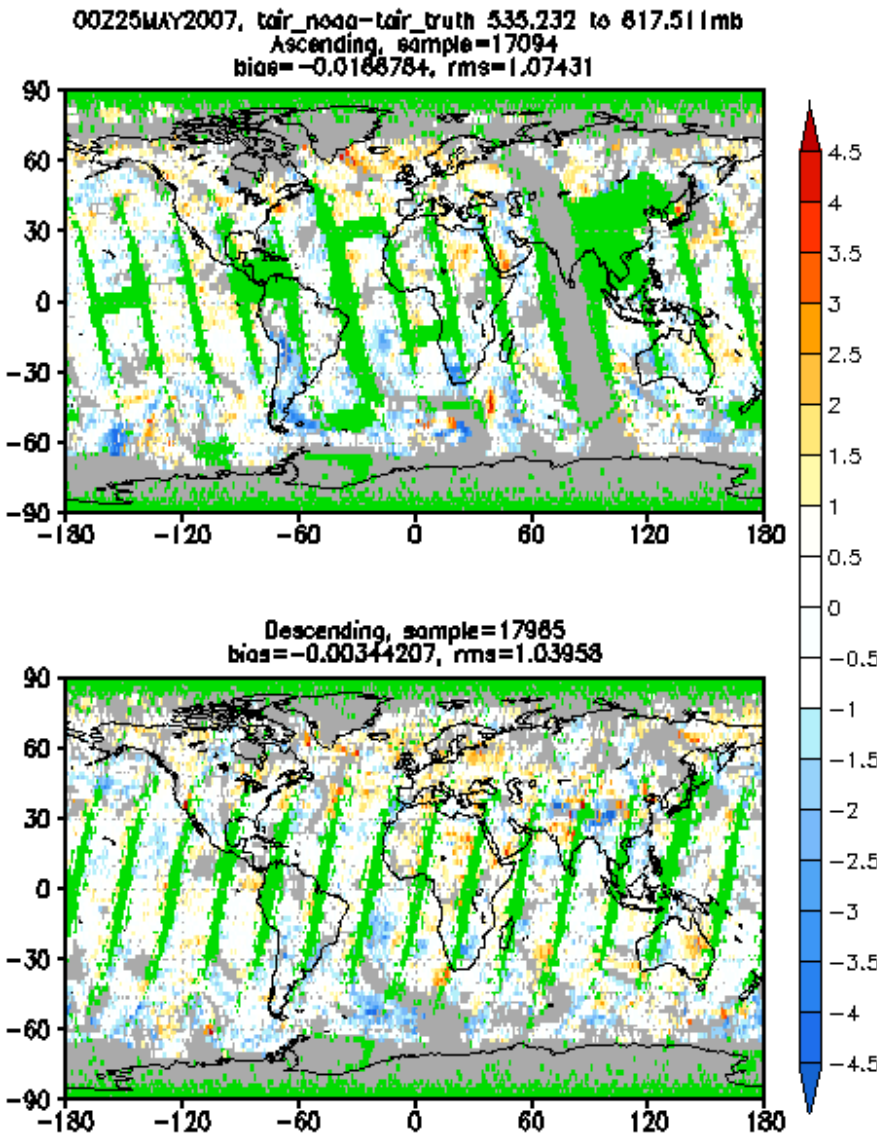


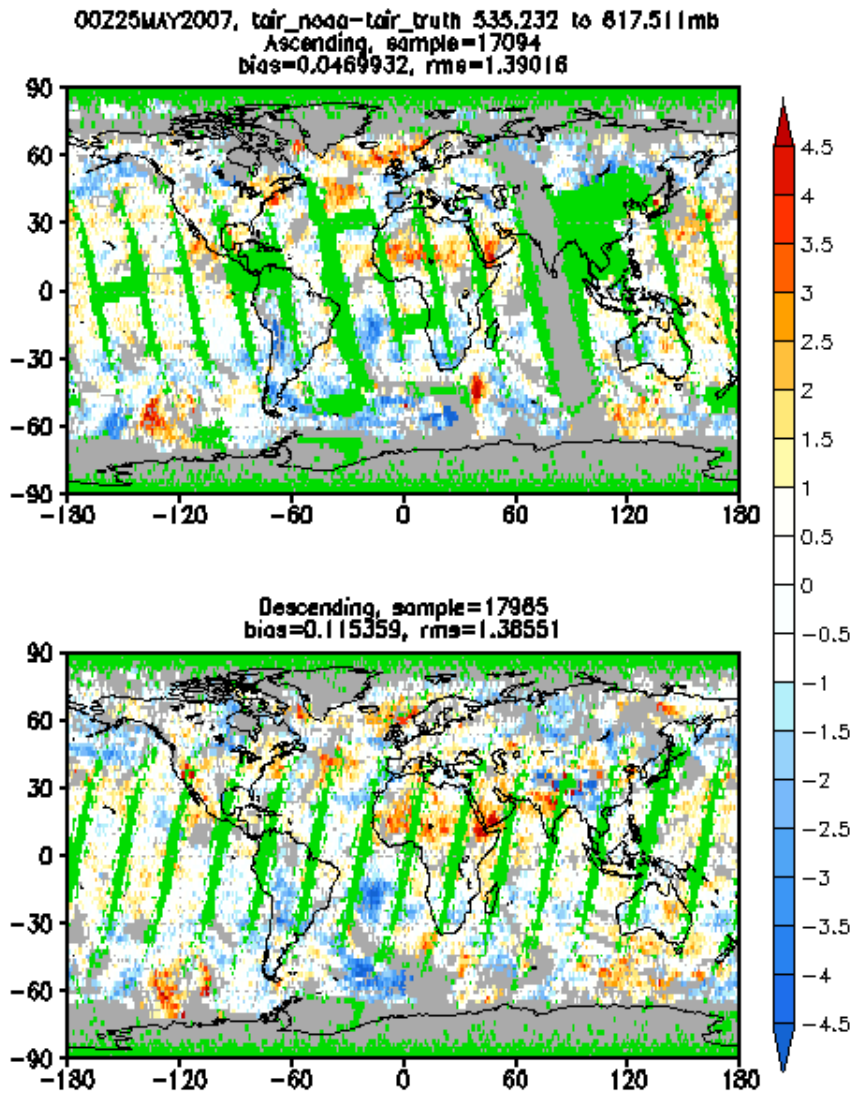
Predictors: AMSU + IASI



Predictors: AMSU

Regression-ECMWF: Temperature at 535.232 — 617.511 hPa

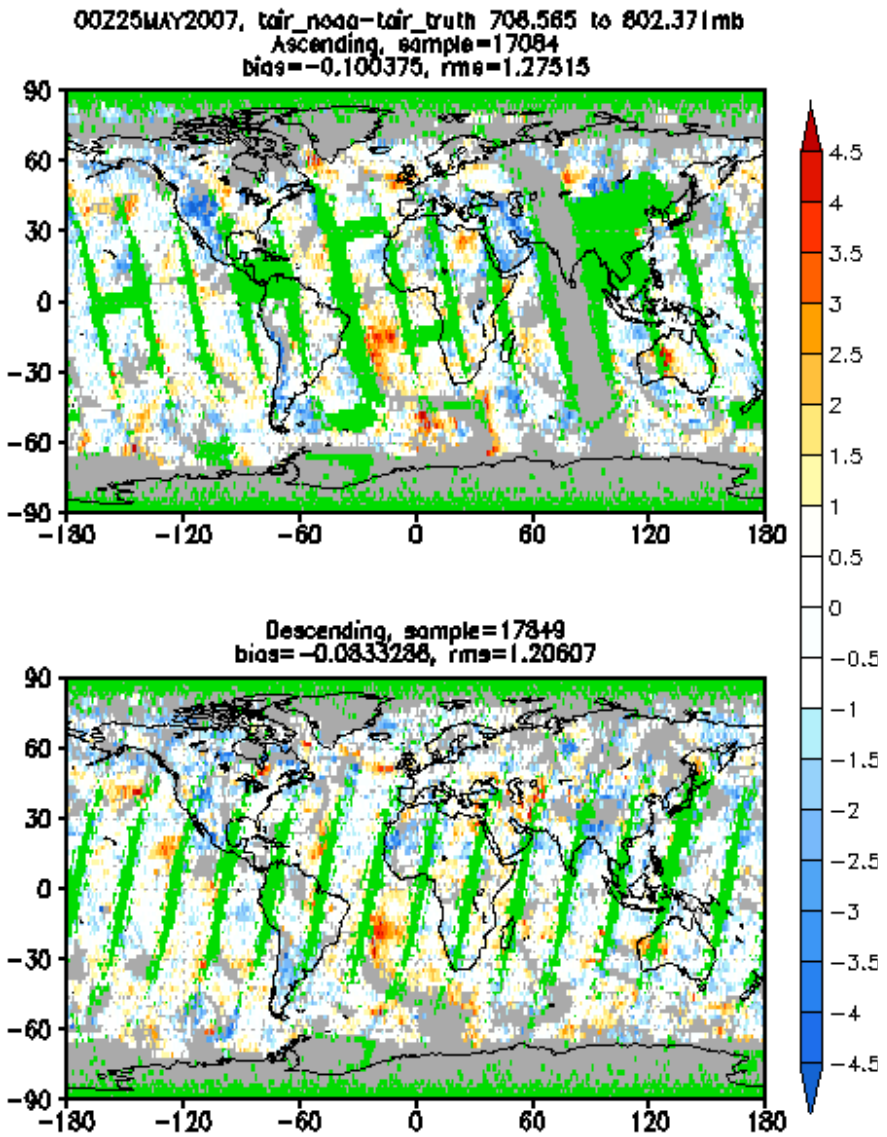




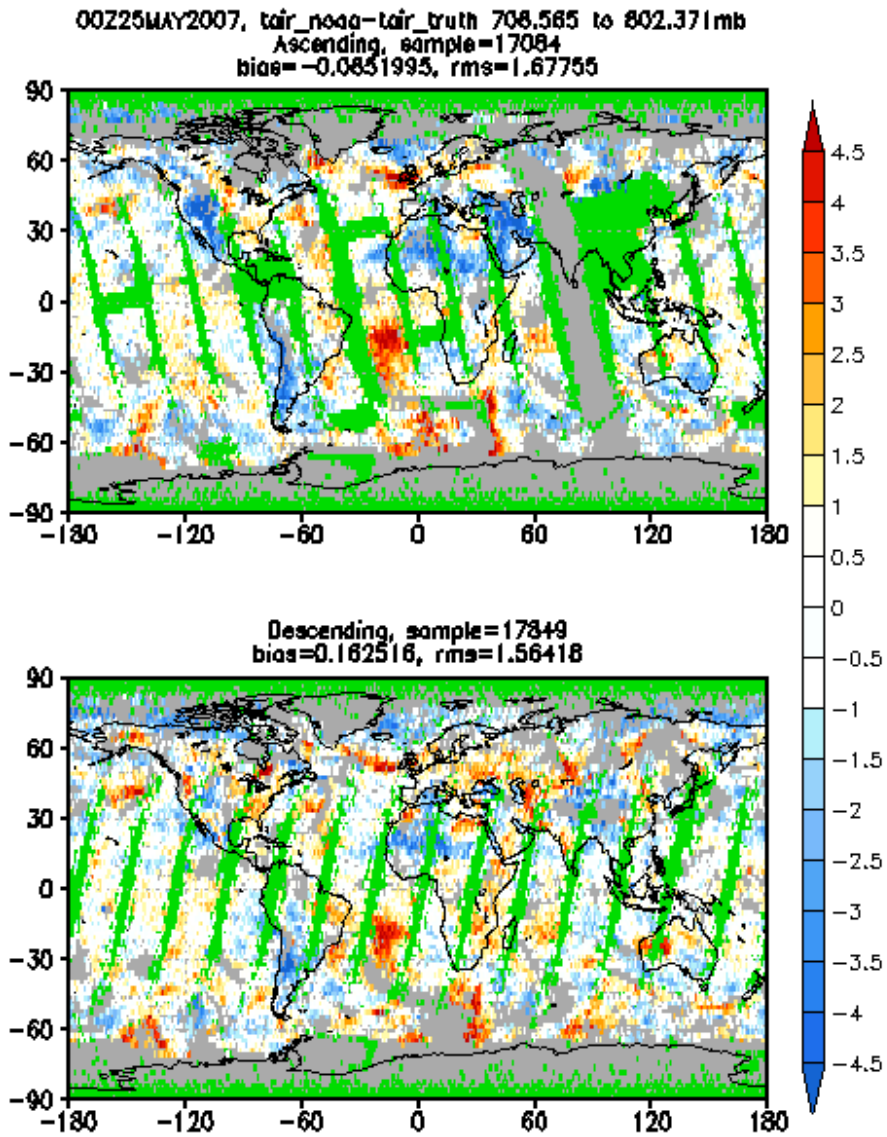
Predictors: AMSU + IASI

Predictors: AMSU

Regression-ECMWF: Temperature at 706.565 — 802.371 hPa

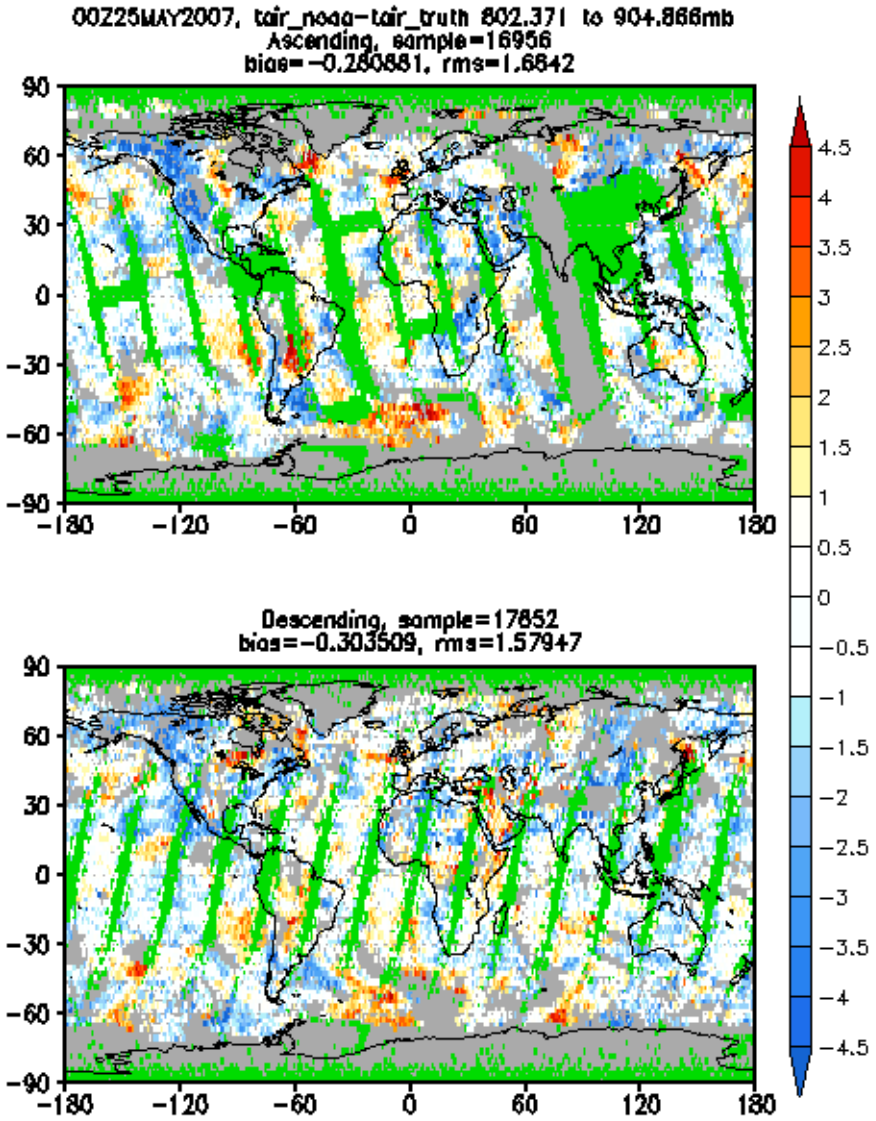


Predictors: AMSU + IASI

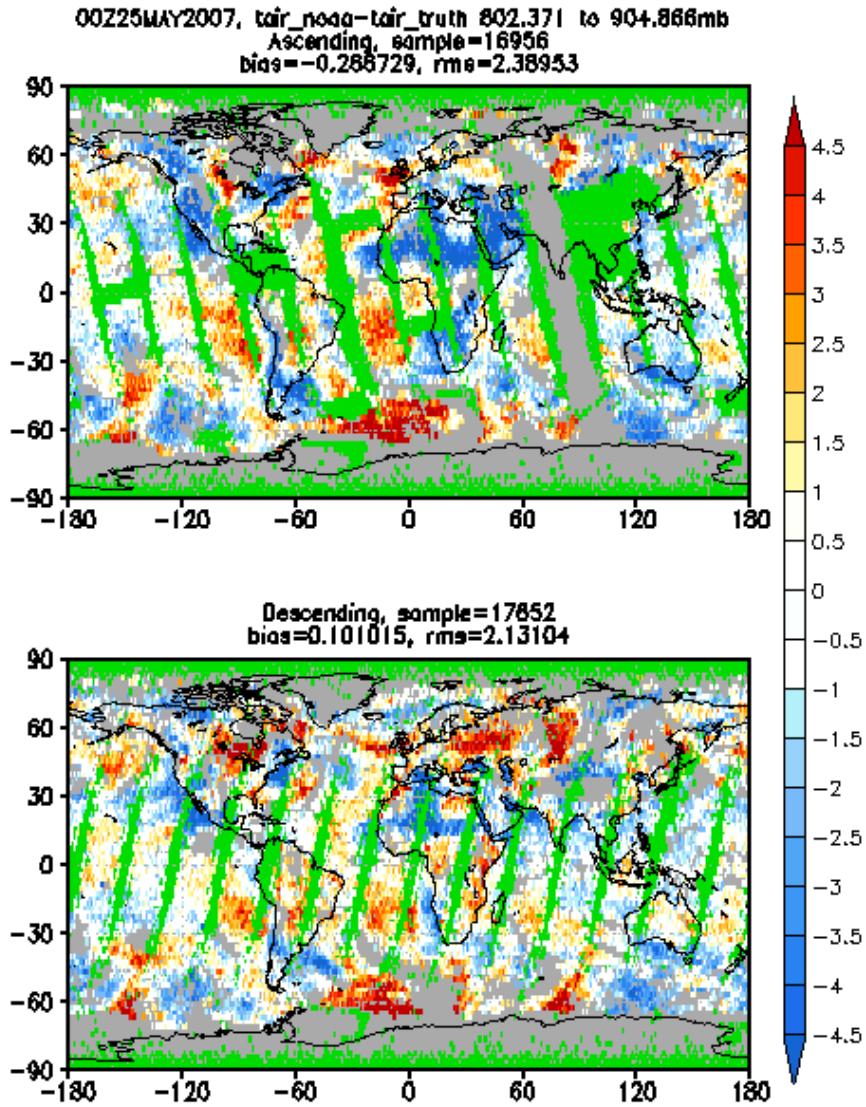


Predictors: AMSU

Regression-ECMWF: Temperature at 802.371 — 904.866 hPa



Predictors: AMSU + IASI



Predictors: AMSU