Principal Component Analysis of AIRS Data

Mitch Goldberg, NOAA/NESDIS

Chris Barnet, Walter Wolf, Lihang Zhou
Murty Divakarla

,

NESDIS Office of Research and Applications

June 30, 2004

ECMWF Workshop

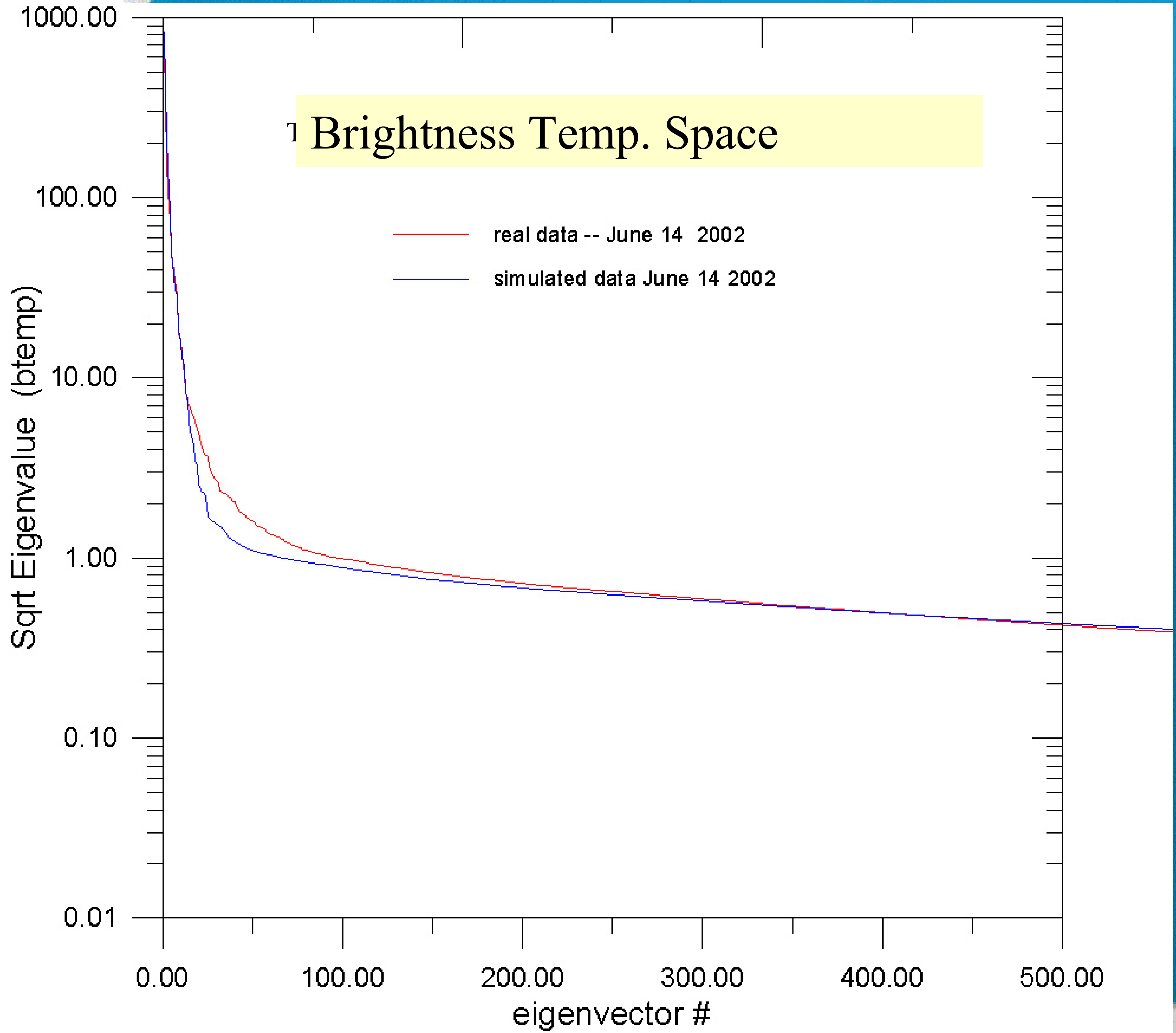
Principal Component Analysis

Regression retrieval

Radiance compression/noise filtering

Case-dependent (dynamic) noise estimation

Quality control



Brightness Temp. Space
real data -- June 14 2002
simulated data June 14 2002
1000.00
100.00
10.00
1.00
0.10
0.01
Sqrt Eigenvalue (btemp)
0.00
100.00
200.00
300.00
400.00
500.00
eigenvector #

LOG

Topics

- *Generation of Eigenvectors*
- *Monitor EOF representation*
- *Reconstructed radiances and noise estimation*
- *Compare static and granule eigenvectors*
- *Cloud Cleared Radiances*
- *New datasets from NESDIS*

Generating AIRS eigenvectors

Each AIRS data vector has 1608 radiance values.

The radiances are normalized by expected instrumental noise (signal to noise)

Compute the covariance matrix S

Compute the eigenvectors E and eigenvalues Λ

$$S = E \Lambda E^T$$

E = matrix of orthonormal eigenvectors (1608x1608)

Λ *= vector of eigenvalues (explained variance)*

Applying AIRS eigenvectors

On independent data – compute principal component scores.

$P \quad = \quad E^T \; R \quad ; \quad$ elements of $R = (r_i - r_i)\,/n_i$

Invert equation and compute reconstructed radiances R^*.

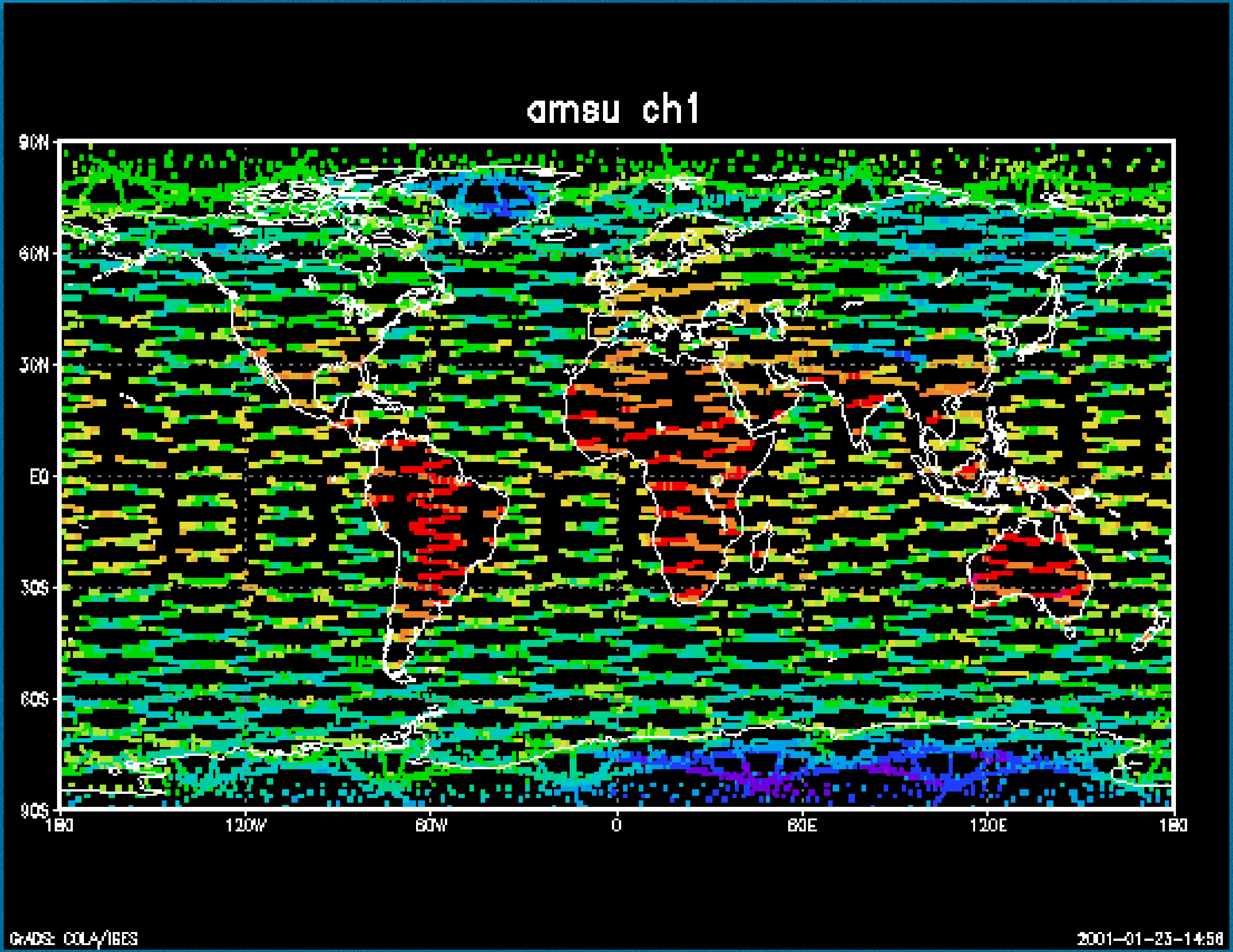
$R^* \quad = \quad E\,P$

Reconstructed radiances are used for quality control.

Reconstruction score $= [\,1/N \;\; 3(R^*_i - R_i)^2\,]^{1/2}$

$i = 1 \ldots.N$ channels

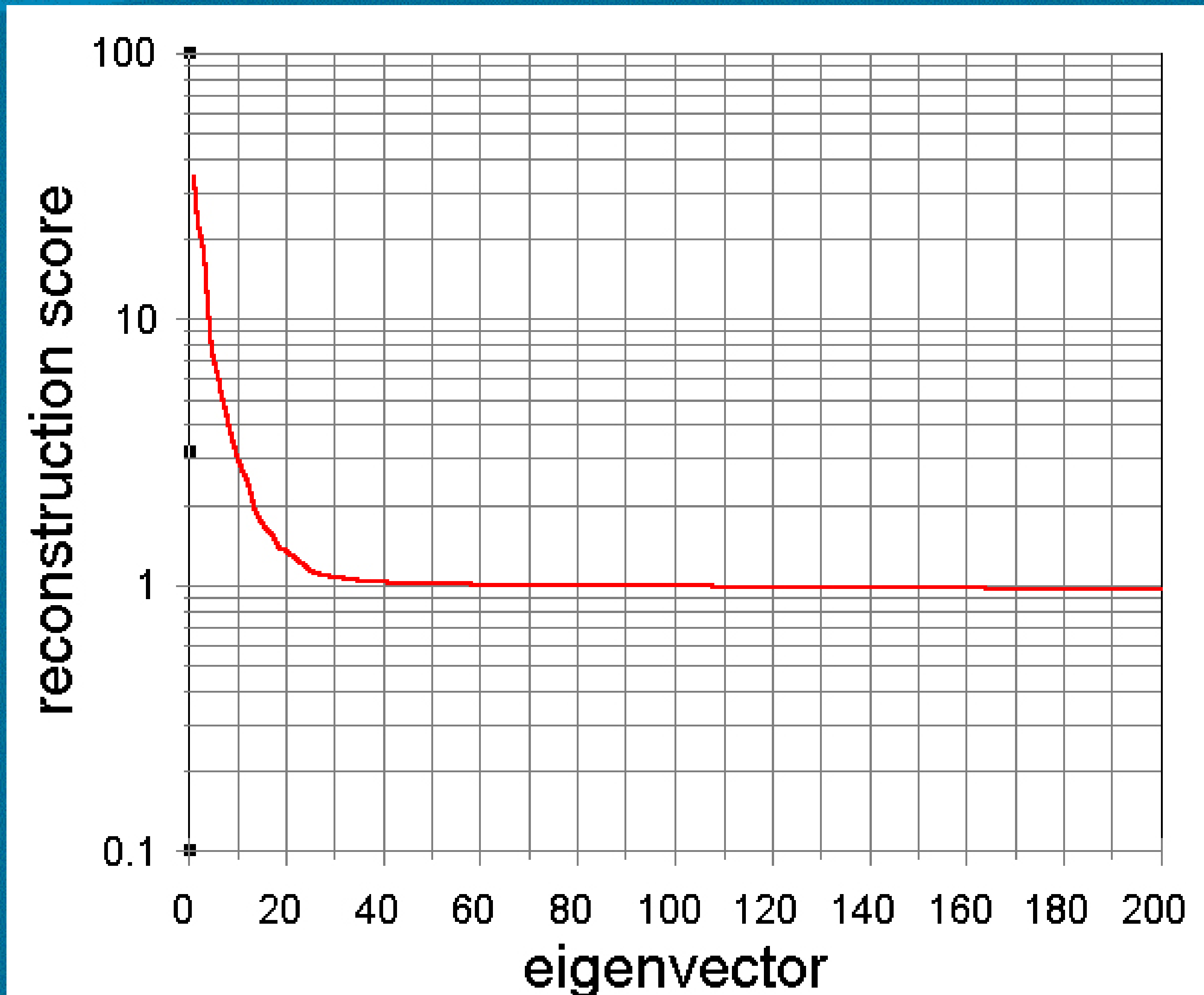
Locations used in generating eigenvectors



Square root of the eigenvalues

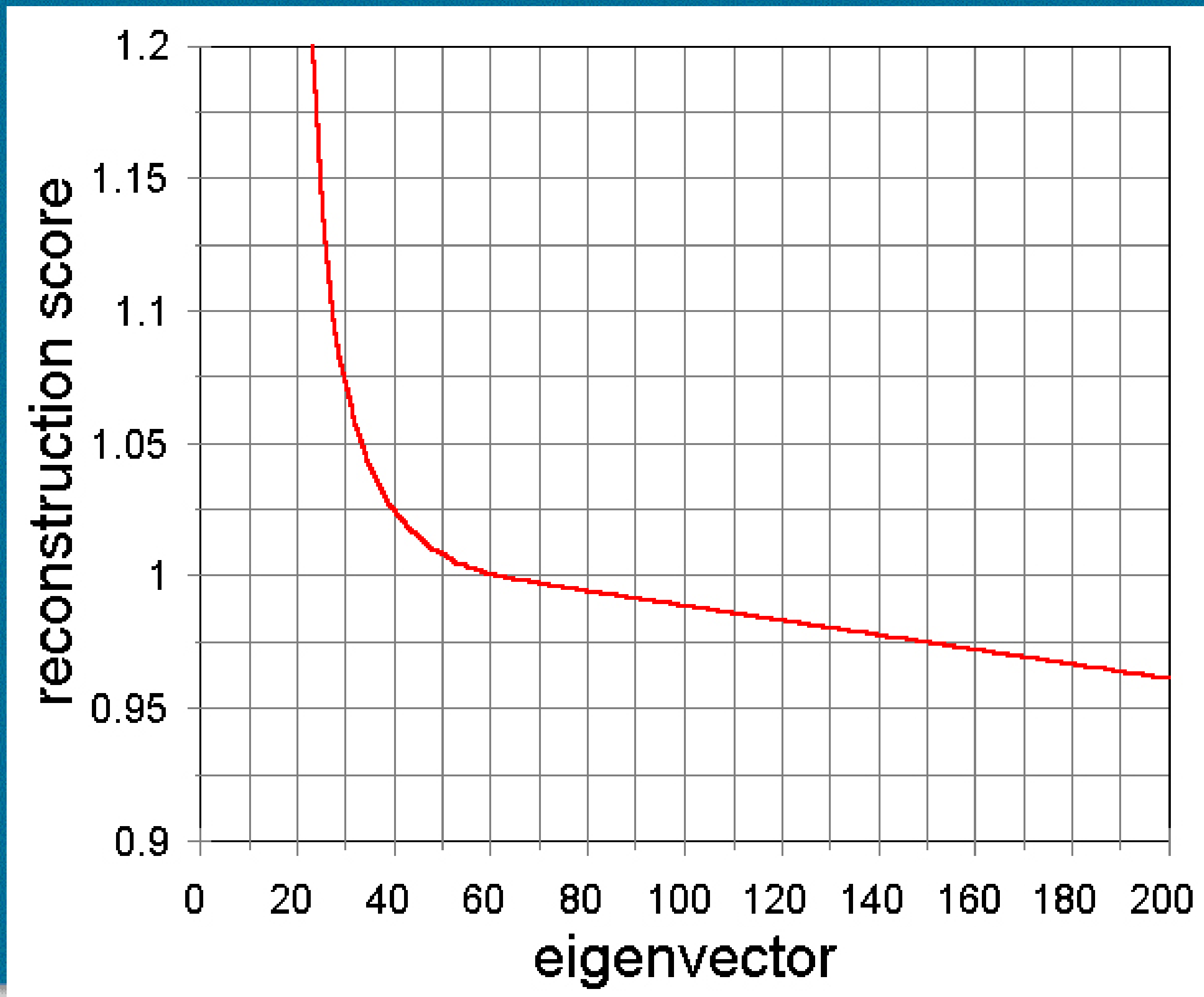
1	7497.60	19	14.68	37	3.38	55	1.25
2	1670.40	20	13.49	38	3.11	56	1.19
3	945.52	21	12.28	39	2.82	57	1.16
4	496.01	22	11.32	40	2.53	58	1.15
5	284.01	23	10.70	41	2.41	59	1.09
6	266.30	24	9.08	42	2.39	60	1.05
7	156.95	25	8.24	43	2.34	61	1.02
8	139.67	26	7.85	44	2.24	62	0.98
9	88.27	27	6.77	45	2.03	63	0.90
10	72.83	28	5.98	46	1.86	64	0.86
11	60.03	29	5.83	47	1.78	65	0.81
12	53.42	30	5.39	48	1.71	66	0.80
13	45.01	31	5.34	49	1.65	67	0.78
14	39.72	32	4.98	50	1.61	68	0.77
15	34.54	33	4.34	51	1.54	69	0.73
16	26.57	34	4.09	52	1.52	70	0.72
17	22.62	35	3.62	53	1.35	71	0.70
18	17.60	36	3.48	54	1.34	72	0.66

- Reconstruction score = $[\ 1/N\ \ 3(R^*_i - R_i)^2\ ]^{1/2}$
  $i = 1\ \ldots.N$ channels

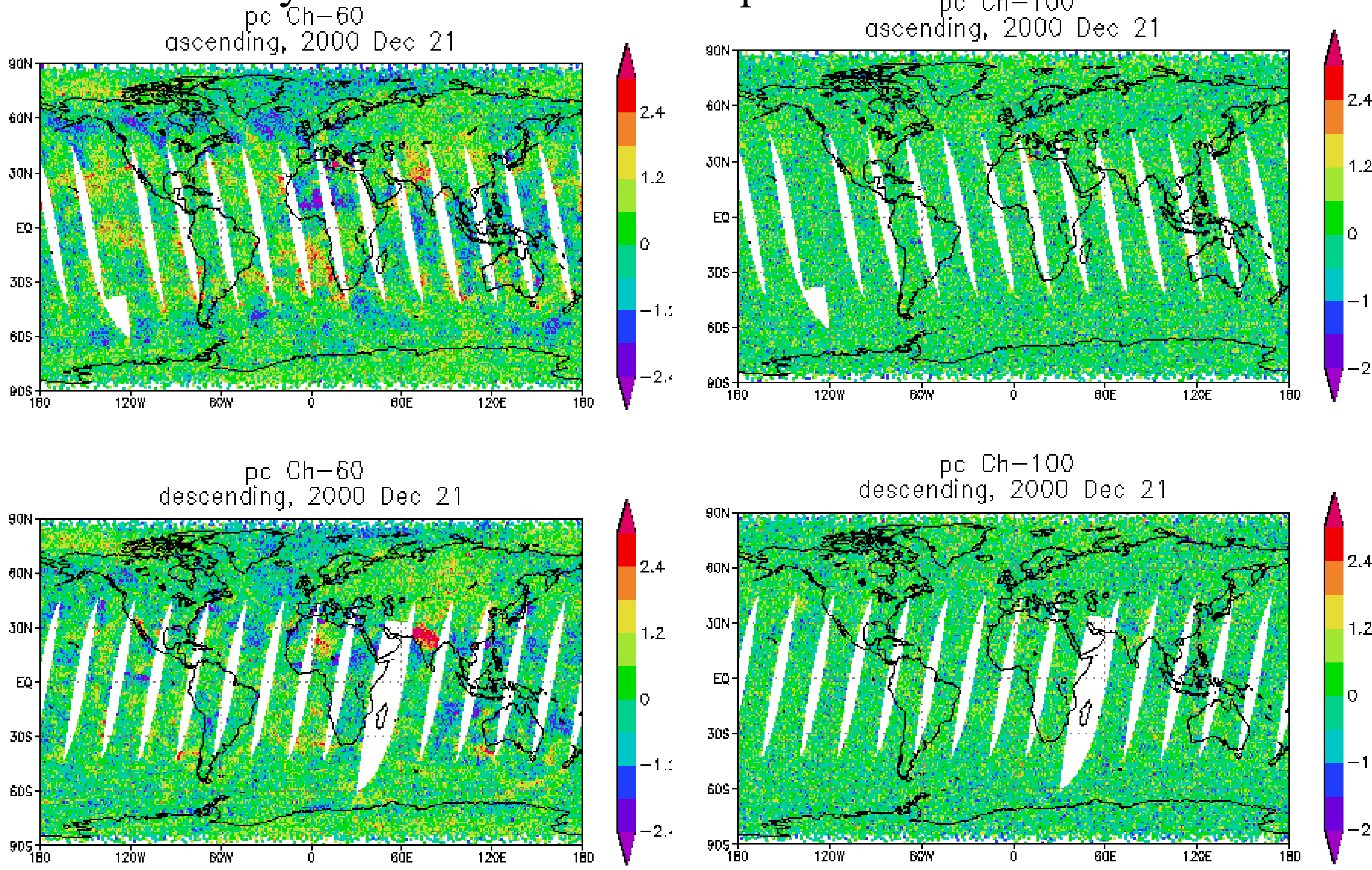


- Reconstruction score $= [\ 1/N\ \ 3(R^*_i - R_i)^2\ ]^{1/2}$

$i = 1 \ldots .N$ channels



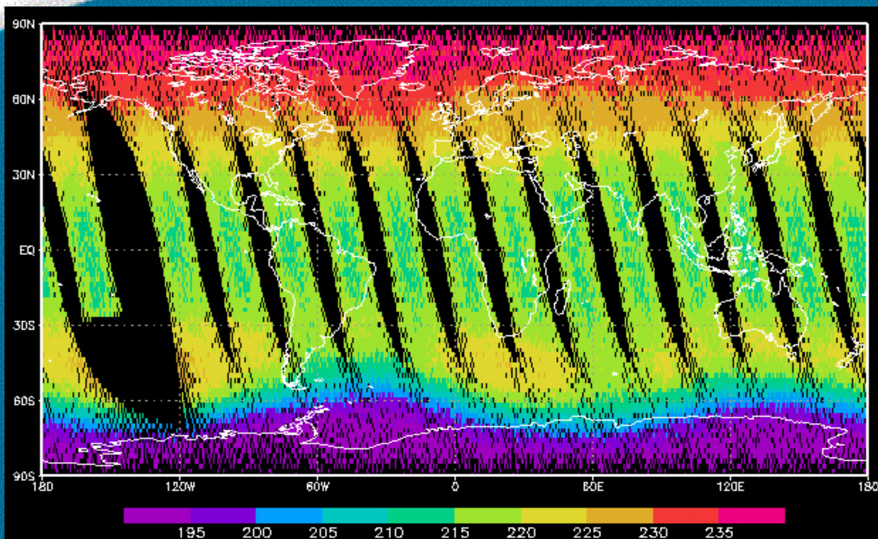
Beware of early truncation -- look at spatial correlations of PC scores



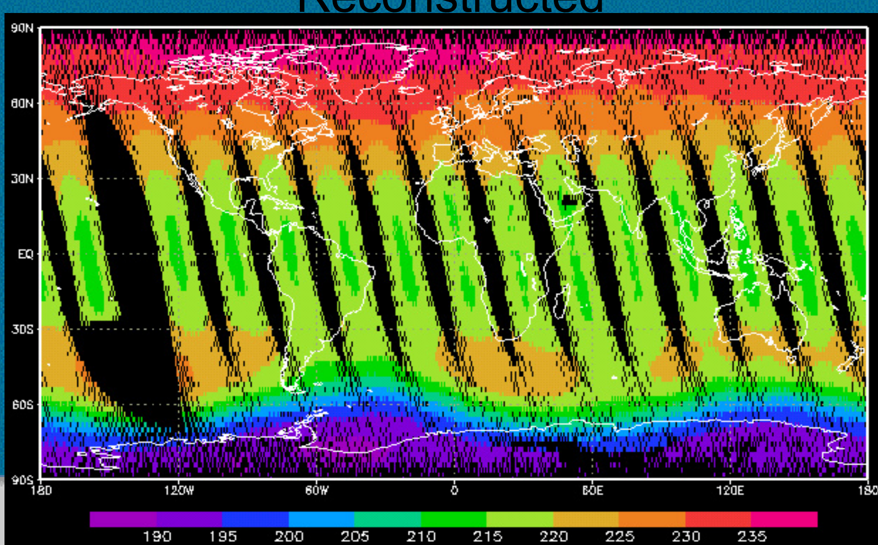
Example of Reconstructed Radiances

649.48 cm-1

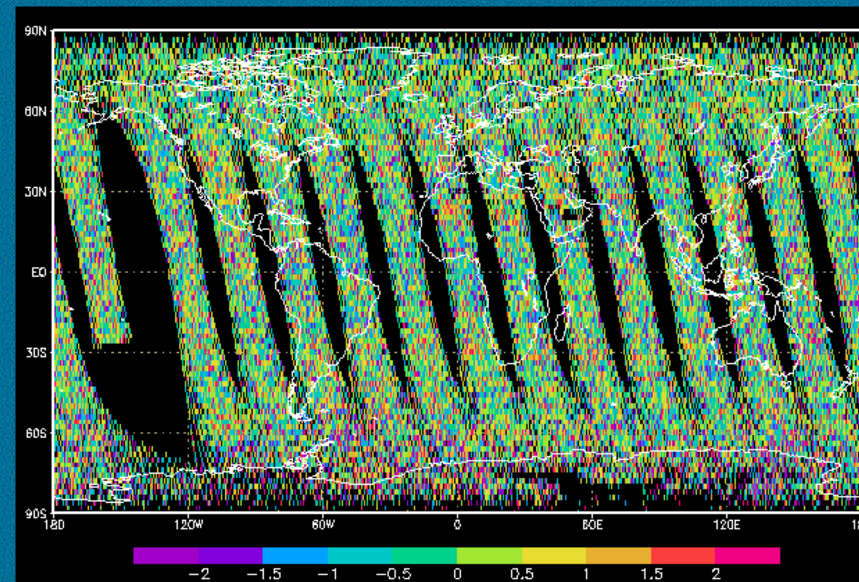
Observed

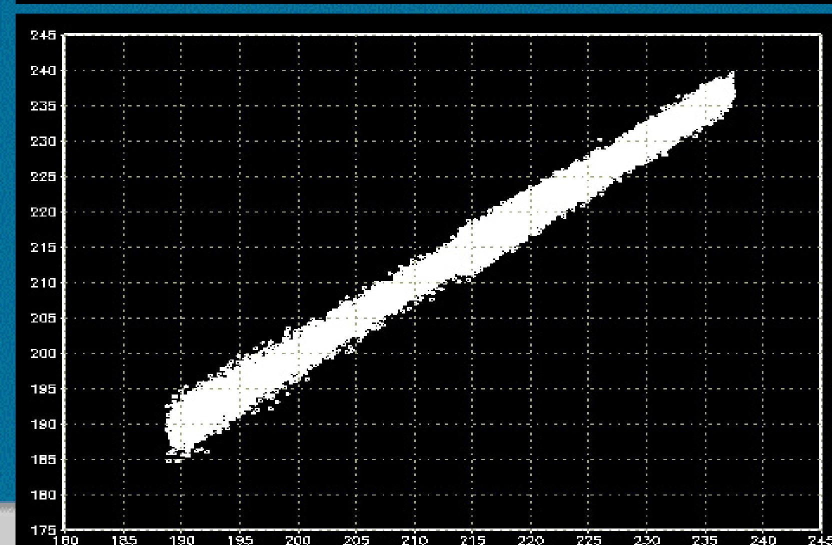


Reconstructed



Reconstructed minus Observed

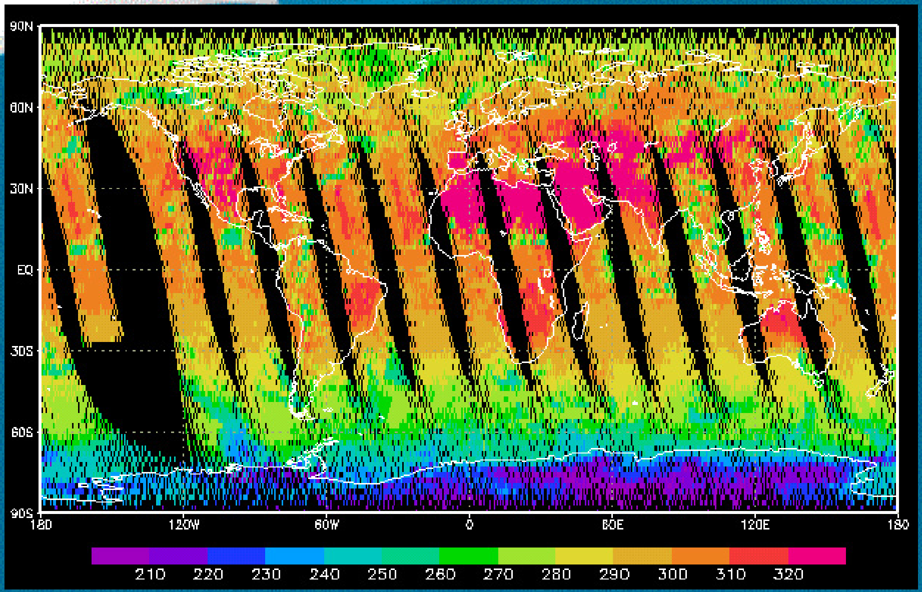




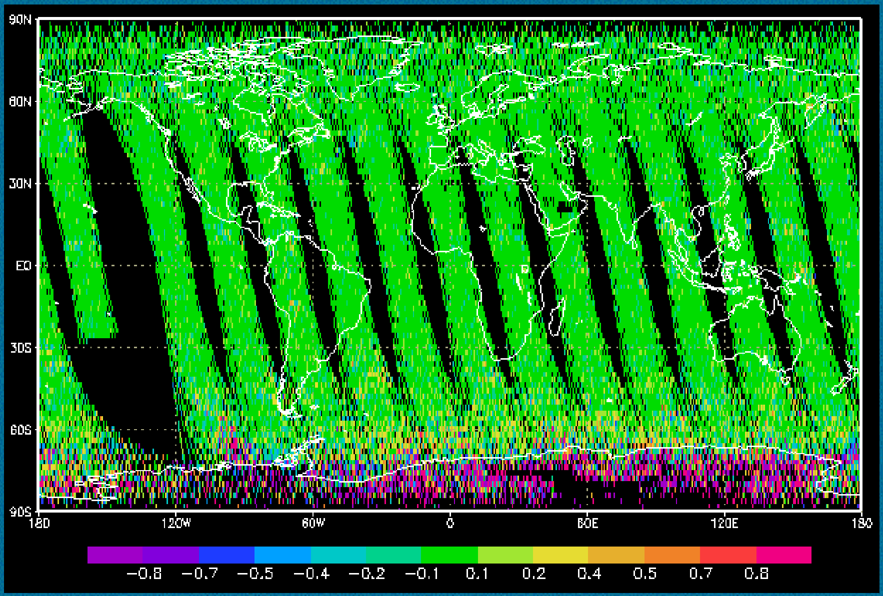
Example of Reconstructed Radiances

2616 cm-1

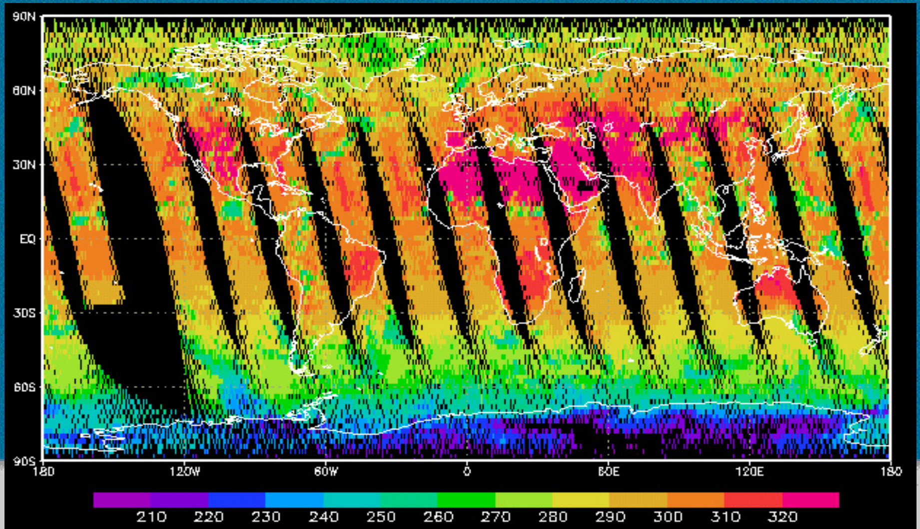
Observed

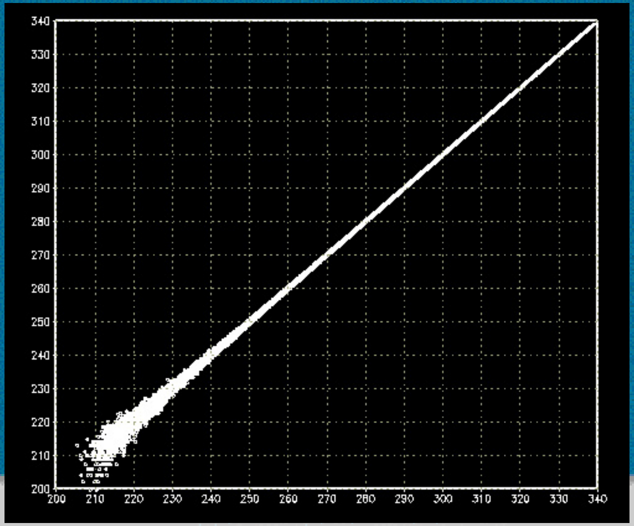


Reconstructed minus Observed

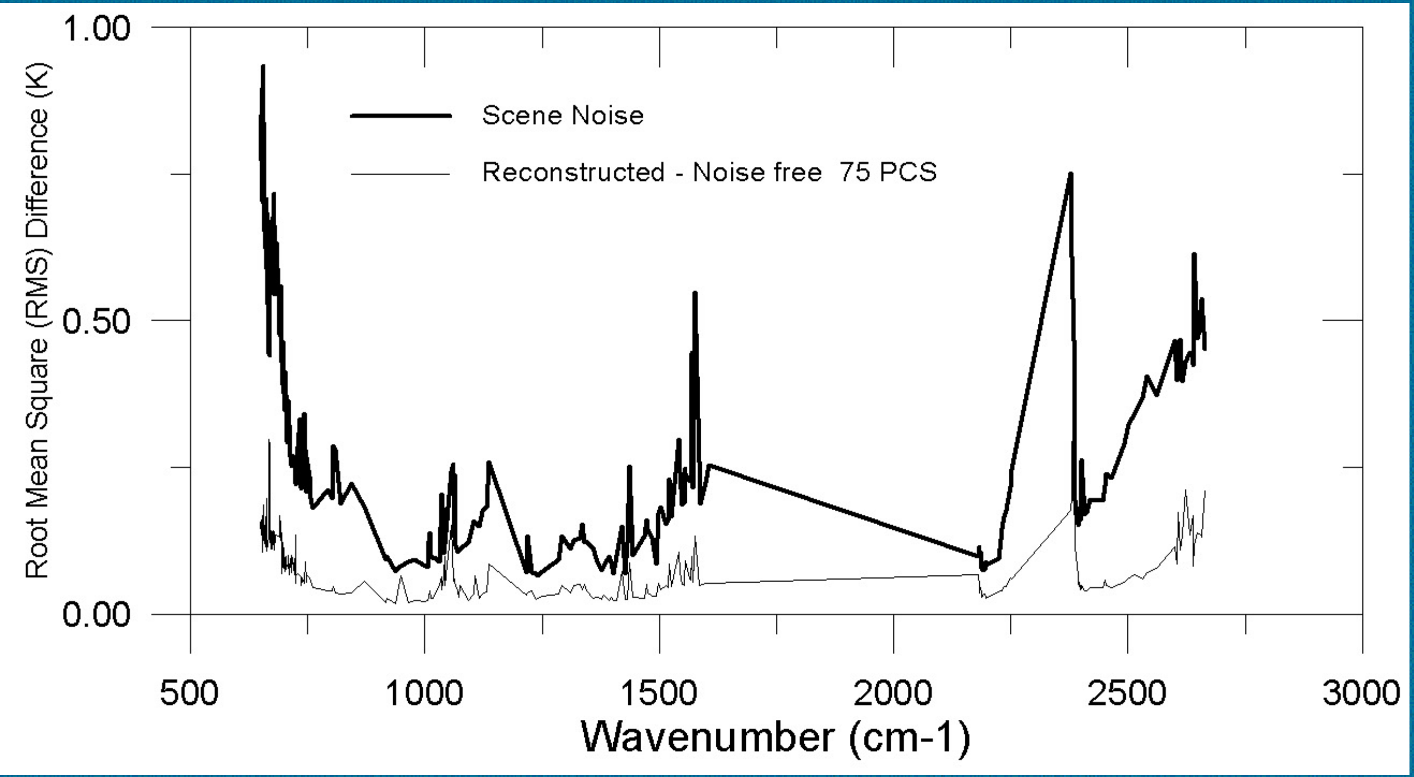


Reconstructed



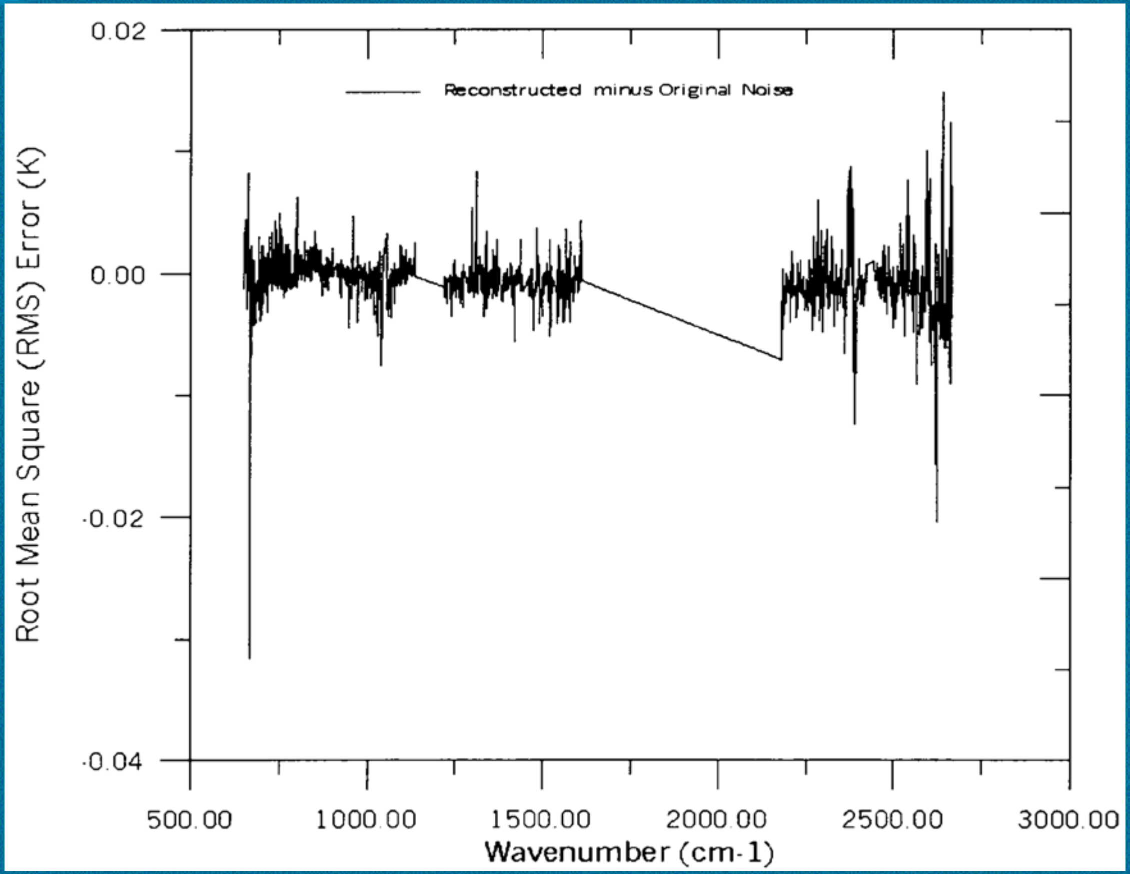


AIRS Instrument Noise vs. Reconstructed Radiance Noise



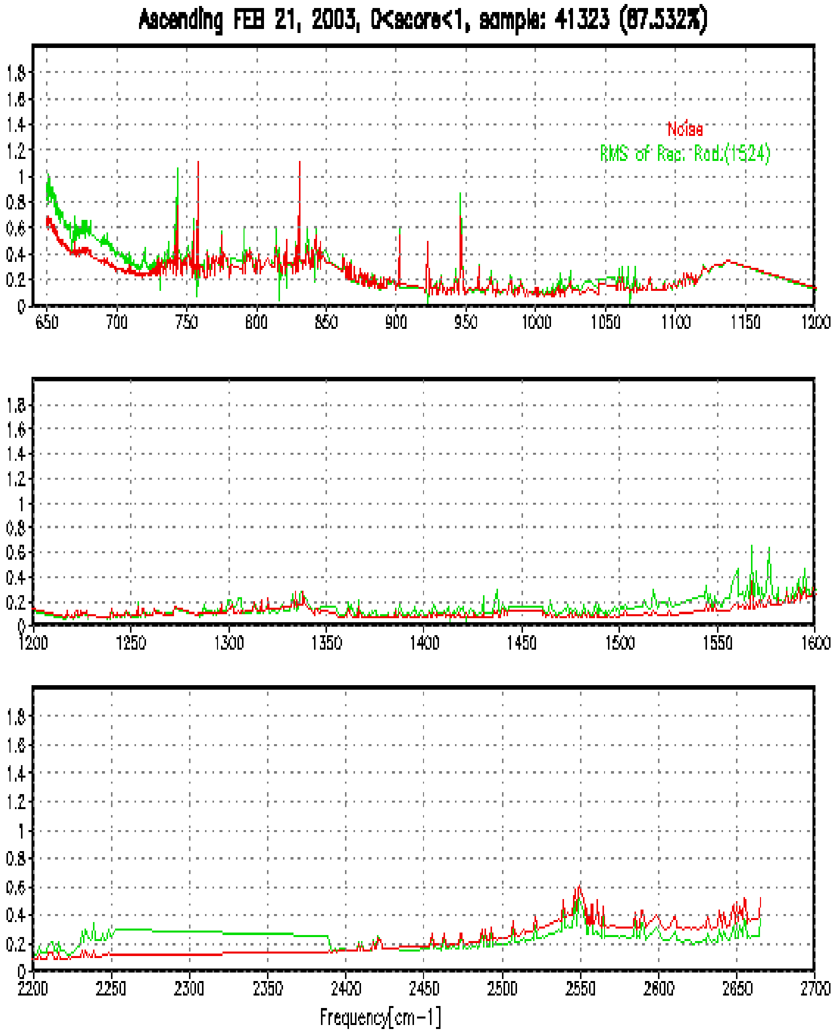
Simulation

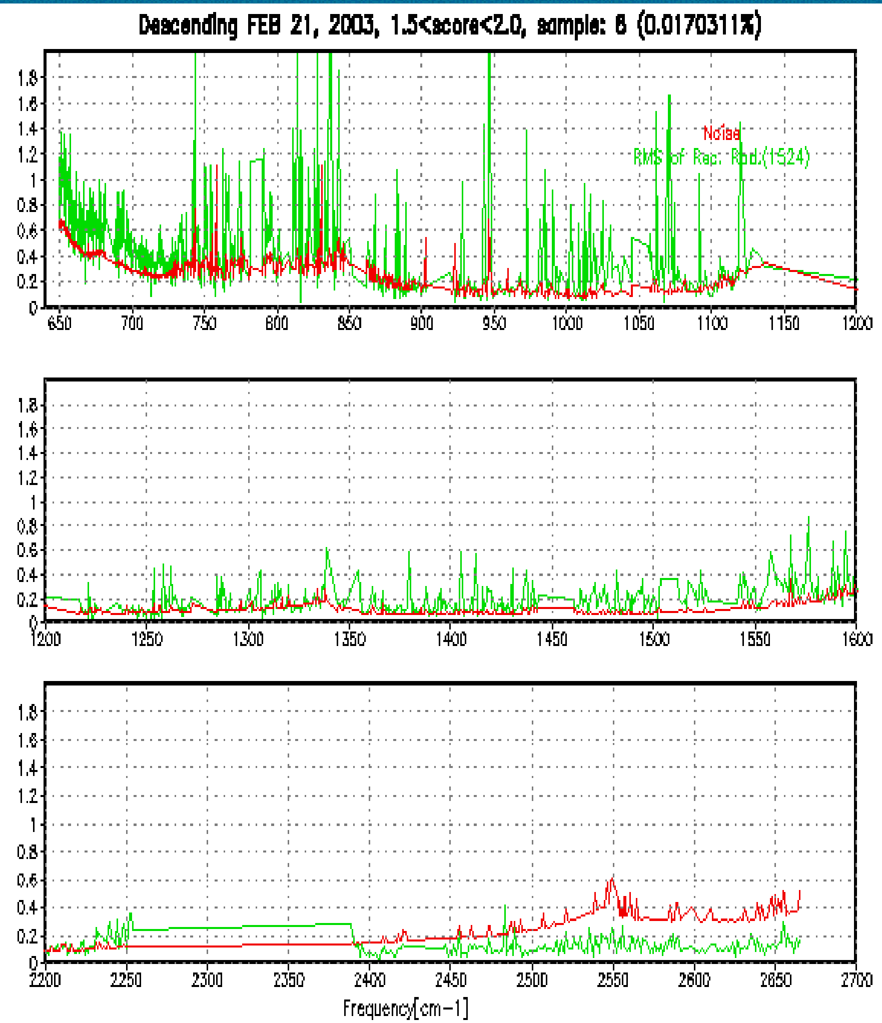
Noise RMS minus Estimated Noise RMS



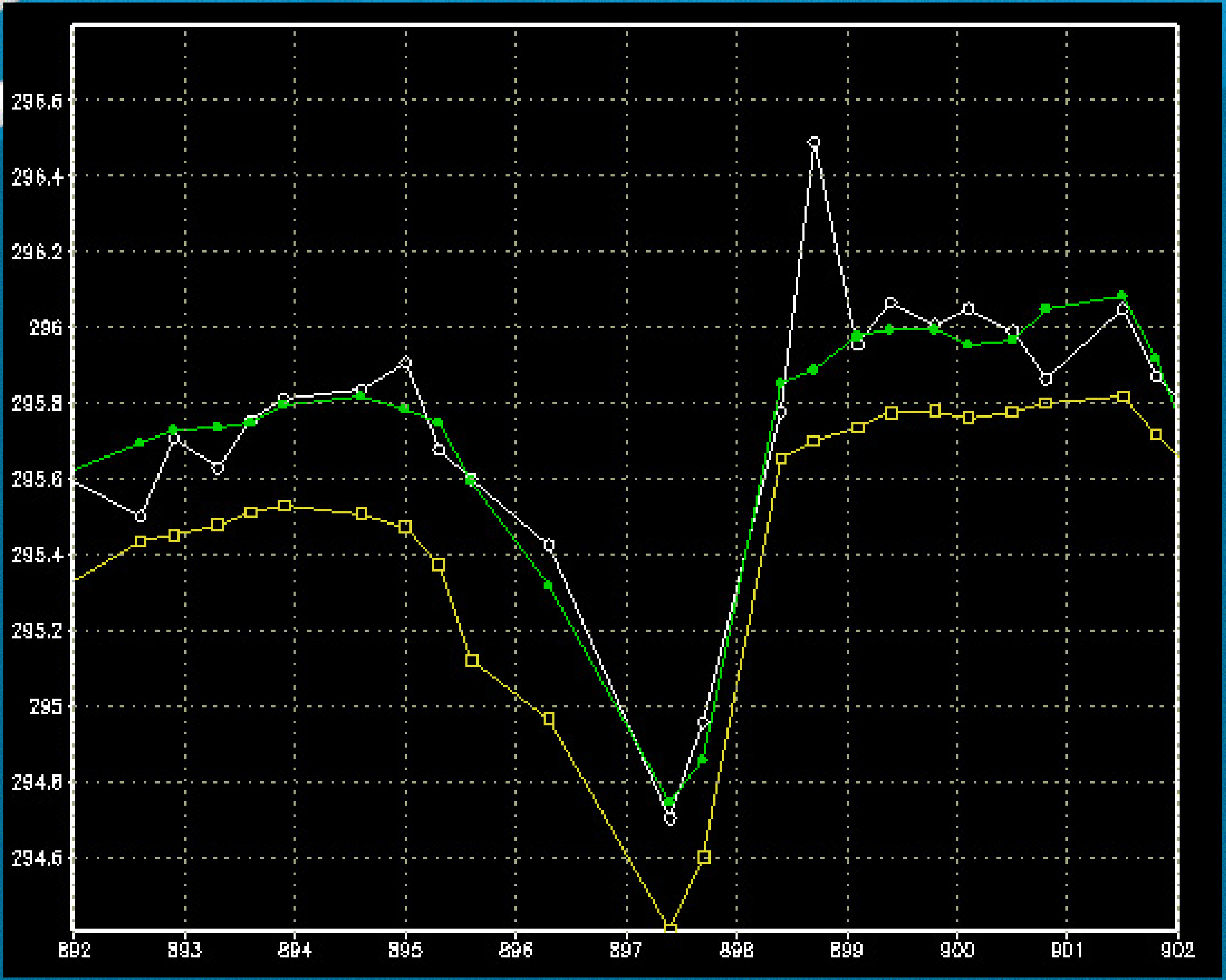
Simulation

Example of QC using PCA reconstructed radiances

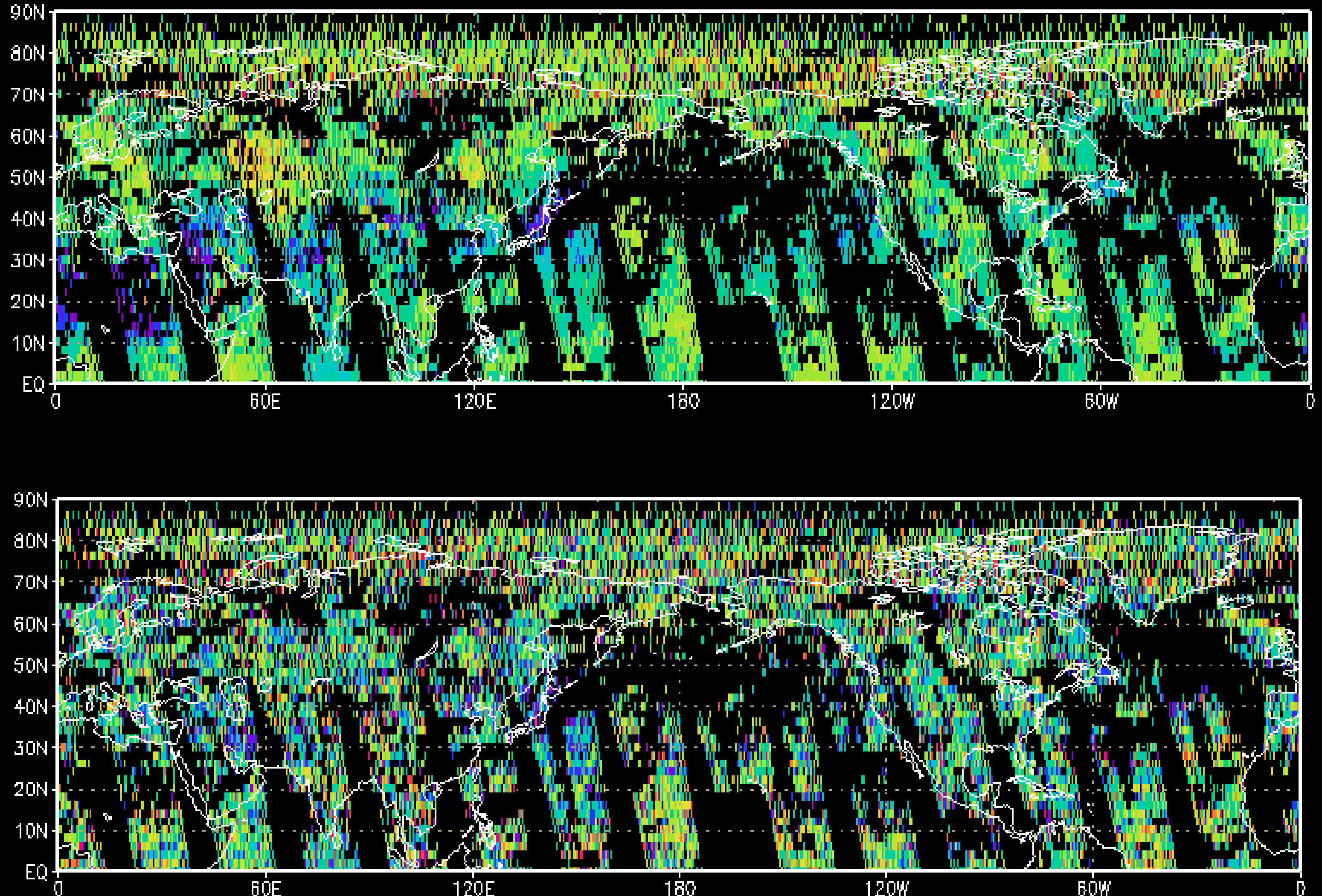




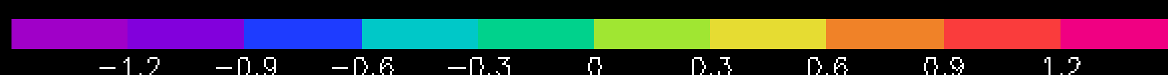
Spectrum from 892 to 902 wavenumber



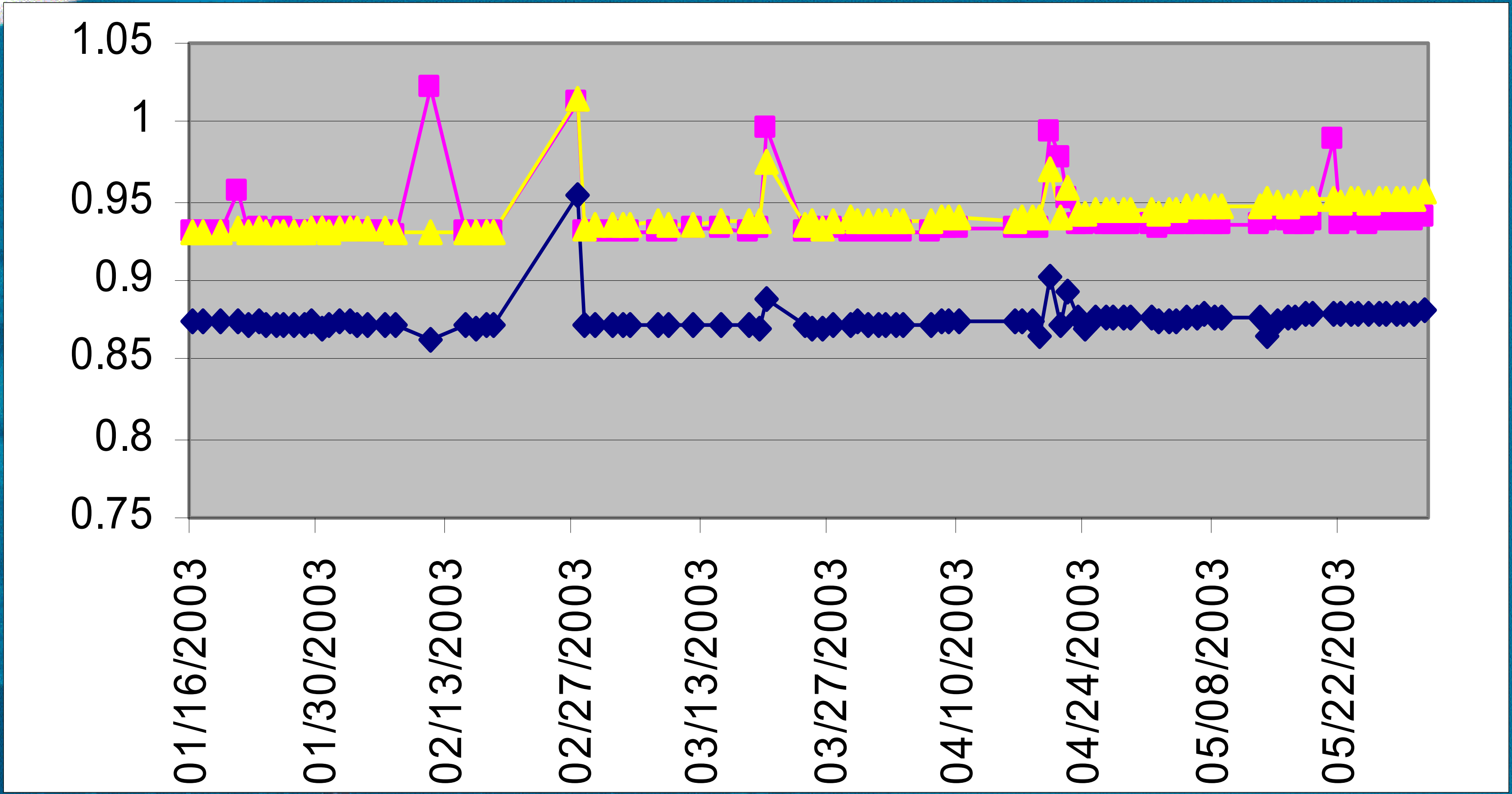
Green is reconstructed, white is original, yellow is calculated from model



original - ecmwf ^^^^^^ rms = 0.92K



Monitoring EOF representation
- dependent (blue), static (yellow), rolling (magenta)



AIRS NEAR REAL TIME Processing System - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Search Favorites Media

Address http://orbit35i.nesdis.noaa.gov/crad2/airsgrid/AIRS_NRT/

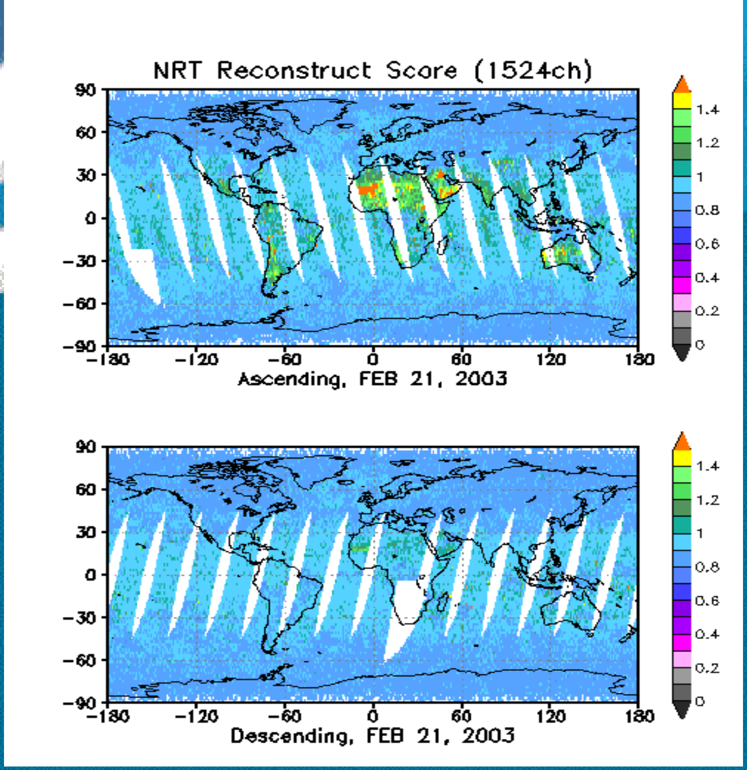
- About AQUA
- AQUA News
- AQUA Orbit
- Global Display
- L1B Radiance
- L2 Retrieval
- EOF Scores
- Quick Browse
- Reconstruct RMS
- Focus Day
- Clear Study 1
- Clear Study 2
- L2 Retrieval Study
- AVN vs ECMWF
- Animation of Today
- AMSU (~1.5 MB)
- HSB (~0.5 MB)
- AIRS (~30 MB)
- EOF (~20 MB)
- Further Info.
- EDOS Transfer Time
- AIRS 281 Ch List
- AIRS 324 Ch. List
- Interface for NRT
- List of Products
- Granule Maps

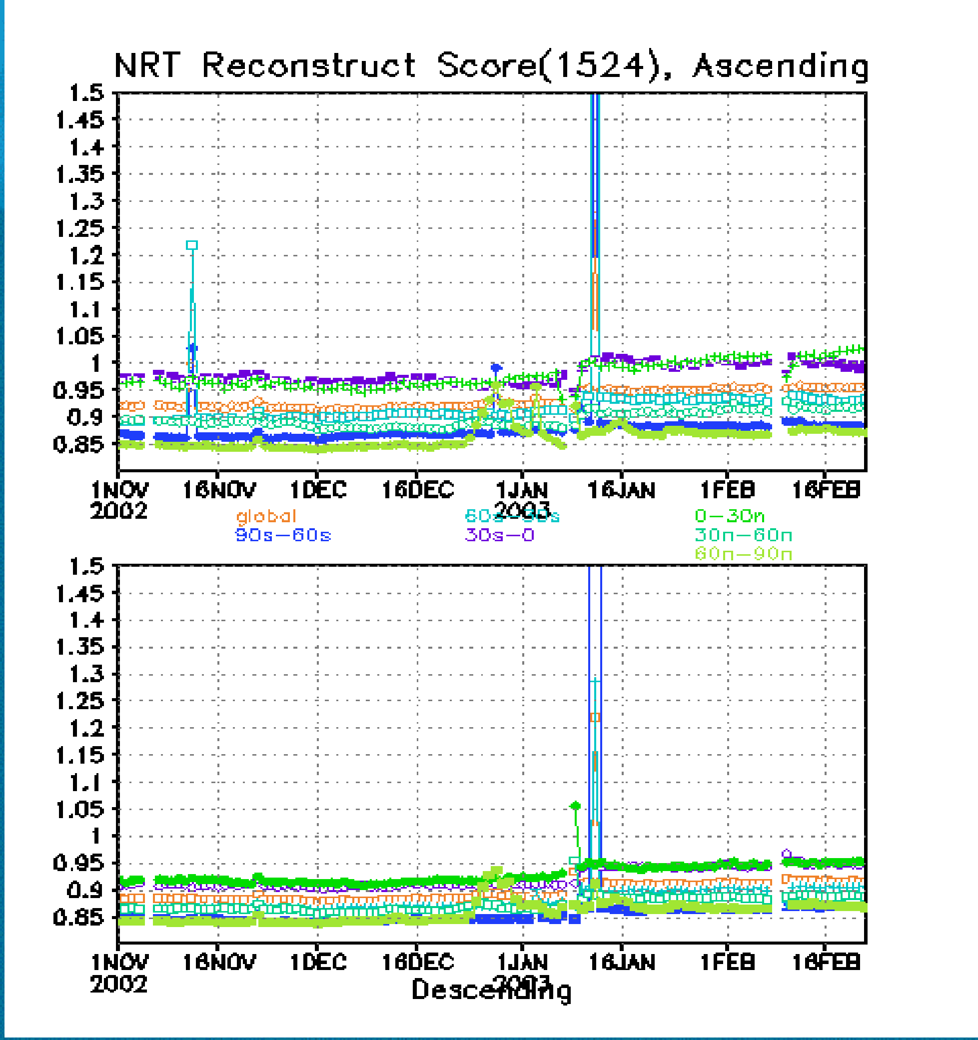
Team Links Go

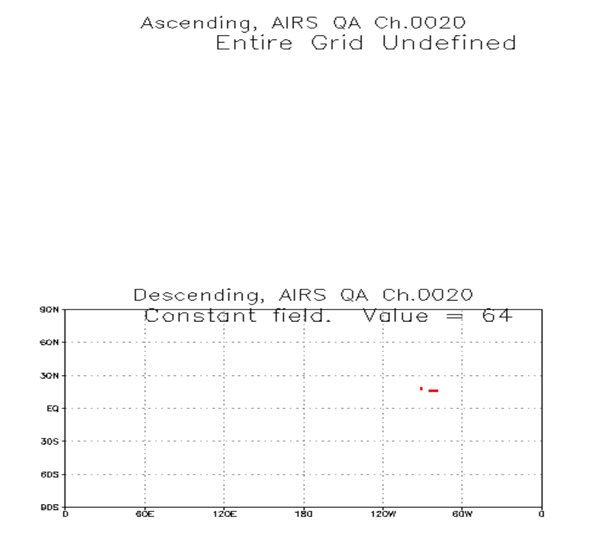
interactive display | Score Map | **RMS of Reconstruct Radiance** | Time Series | qa>0 | rms>5*noise

ch.	freq.	noise	obs/recon.	s0-1 asc.	s0-1 desc.	s1-1.5 asc.	s1-1.5 desc.	s1.5-2 asc.	s1.5-2 desc.	s>2 asc.	s>2 desc.
		sample		41323	44794	5699	2085	172	8	10	1
		percent		87.532	95.361	12.072	4.439	0.364	0.017	0.021	0.002
1	649.613	0.637	216.54/ 216.62	0.881	0.888	0.942	0.944	0.959	1.173	1.008	0.014
2	649.851	0.679	214.62/ 214.90	0.949	0.959	1.049	1.041	1.289	1.160	1.859	1.566
3	650.090	0.684	217.84/ 217.74	0.941	0.948	1.015	0.988	1.242	0.915	1.172	0.311
4	650.328	0.640	227.27/ 227.41	0.815	0.821	0.854	0.843	0.906	0.639	0.584	1.461
5	650.567	0.598	226.06/ 226.02	0.786	0.783	0.797	0.829	0.932	0.940	0.767	0.037
6	650.807	0.649	217.42/ 217.43	0.892	0.893	0.962	0.946	1.018	0.812	1.526	1.086
7	651.046	0.677	211.77/ 211.91	0.982	0.982	1.080	1.075	1.134	0.907	0.928	0.765
8	651.285	0.697	210.87/ 210.54	1.009	1.017	1.139	1.121	1.287	1.374	1.080	1.297
9	651.525	0.662	211.98/ 211.92	0.943	0.955	1.082	1.043	1.365	0.972	1.142	0.815
10	651.765	0.641	223.07/ 223.12	0.846	0.851	0.908	0.896	1.016	1.099	1.195	0.325
11	652.005	0.670	227.73/ 227.61	0.864	0.866	0.886	0.857	0.937	0.945	0.740	0.670
12	652.727	0.616	211.07/ 211.35	0.889	0.907	1.006	1.020	0.912	0.978	1.025	2.059
13	652.968	0.633	210.74/ 210.82	0.897	0.913	1.010	1.048	1.142	1.352	1.369	0.421
14	653.209	0.667	219.06/ 219.01	0.911	0.913	0.969	0.985	0.984	0.716	0.955	0.452
15	653.450	0.624	227.97/ 228.12	0.808	0.815	0.822	0.803	1.138	1.137	1.419	0.755
16	653.691	0.617	221.66/ 221.71	0.828	0.833	0.881	0.853	0.967	1.061	1.183	0.063
17	653.933	0.602	214.93/ 215.03	0.846	0.853	0.920	0.930	1.180	0.656	1.377	0.247
18	654.175	0.580	212.07/ 212.01	0.833	0.830	0.955	0.906	0.997	0.555	0.977	0.719
19	654.417	0.625	210.17/ 210.24	0.918	0.928	1.020	1.019	1.115	1.232	1.086	0.012
20	654.659	0.608	215.76/ 215.75	0.850	0.856	0.917	0.930	0.900	1.051	0.657	0.211
21	654.902	0.621	227.69/ 227.82	0.803	0.816	0.829	0.827	0.849	0.854	0.937	1.005
22	655.144	0.581	225.10/ 225.38	0.764	0.768	0.808	0.841	0.913	0.882	0.625	1.735
23	655.387	0.557	216.71/ 216.68	0.775	0.787	0.851	0.846	0.914	0.880	0.817	1.128
24	655.630	0.614	213.97/ 213.91	0.870	0.878	0.984	0.936	1.204	1.040	1.658	0.588
25	655.873	0.576	210.76/ 210.83	0.841	0.843	0.936	0.937	1.010	1.089	0.832	1.002
26	656.116	0.568	212.77/ 213.06	0.812	0.815	0.910	0.878	0.983	0.896	0.810	1.482
27	656.360	0.552	224.81/ 224.80	0.721	0.726	0.779	0.752	0.931	0.795	1.034	0.739
28	656.604	0.539	227.61/ 227.62	0.708	0.705	0.728	0.723	0.933	0.471	1.211	0.088
29	656.848	0.578	218.22/ 218.11	0.799	0.798	0.854	0.830	0.971	1.346	0.913	0.771
30	657.092	0.533	215.14/ 215.14	0.757	0.755	0.845	0.799	1.138	1.283	0.525	0.320
31	657.336	0.514	211.99/ 211.94	0.736	0.747	0.812	0.838	0.917	0.653	1.190	0.052
32	657.581	0.524	211.62/ 211.85	0.750	0.762	0.862	0.837	1.068	0.979	0.784	1.110
33	657.825	0.536	221.43/ 221.42	0.720	0.726	0.770	0.768	1.108	0.842	0.990	0.110
34	658.070	0.517	228.07/ 228.08	0.678	0.675	0.684	0.684	0.792	0.920	0.553	0.049
35	658.315	0.507	220.52/ 220.60	0.687	0.695	0.737	0.736	0.944	0.768	1.099	0.378
36	658.561	0.530	214.81/ 214.71	0.753	0.752	0.824	0.826	0.879	0.487	0.652	0.341

start | Microsoft PowerPoint ... | ret.ctl - WordPad | Print out the Figure - ... | AIRS NEAR REAL TIM... | X Server display:0 | 10:54 AM


NRT Reconstruct Score (1524ch)
Ascending, FEB 21, 2003
Descending, FEB 21, 2003
90
60
30
0
-30
-60
-90
-180
-120
-60
0
60
120
180
1.4
1.2
1
0.8
0.6
0.4
0.2
0


NRT Reconstruct Score(1524), Ascending
1NOV 2002
16NOV
1DEC
16DEC
1JAN 2003
16JAN
1FEB
16FEB
global
90s-60s
60s-30s
30s-0
0-30n
30n-60n
60n-90n
Descending


Ascending, AIRS QA Ch.0020
Entire Grid Undefined
Descending, AIRS QA Ch.0020
Constant field. Value = 64

Static versus Granule Compression Studies

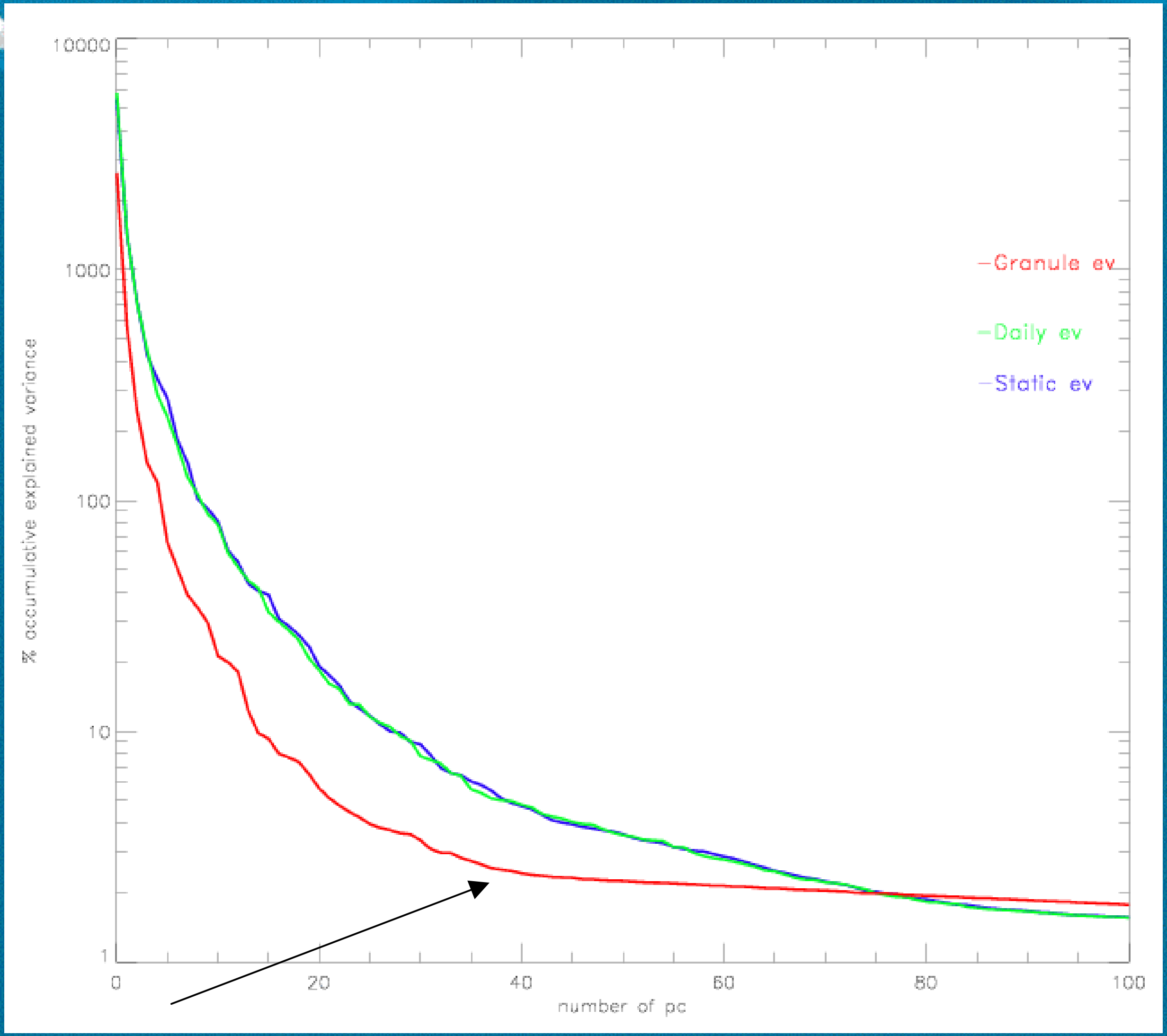
Three sets of eigenvectors were generated

1). Static Eigenvector (from global data of January 15th 2003)

2). Daily Eigenvector (from global data of 2 days earlier from the granule)

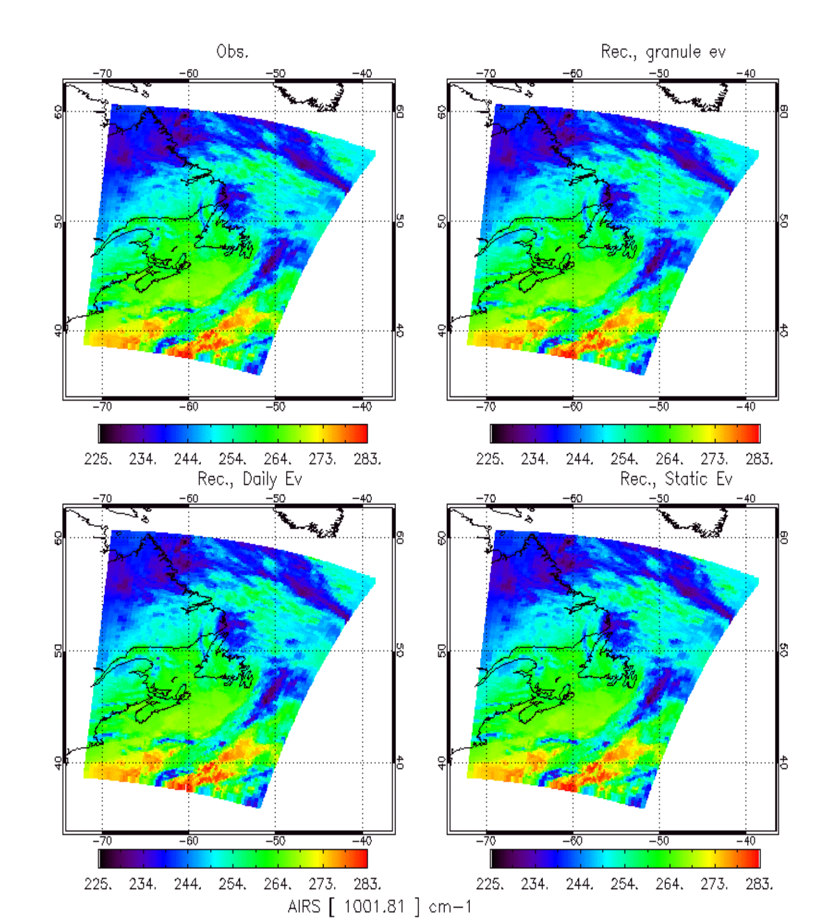
3). Granule Eigenvector (from the data of the current granule)

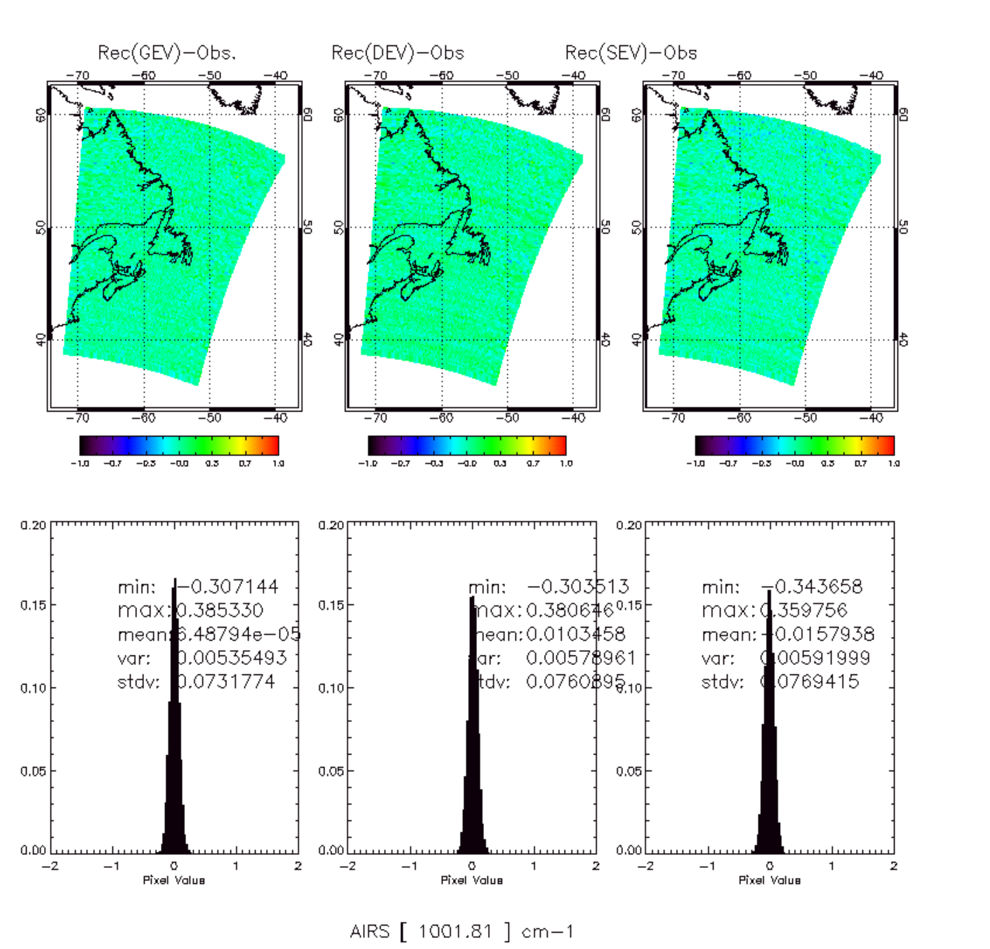
Square Root of Eigenvalue



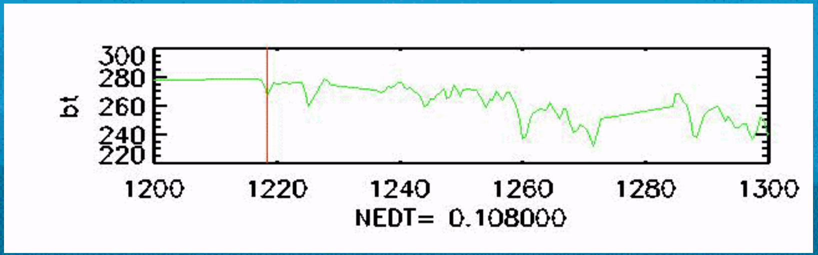
Less PCs are needed for granule-based eigenvectors

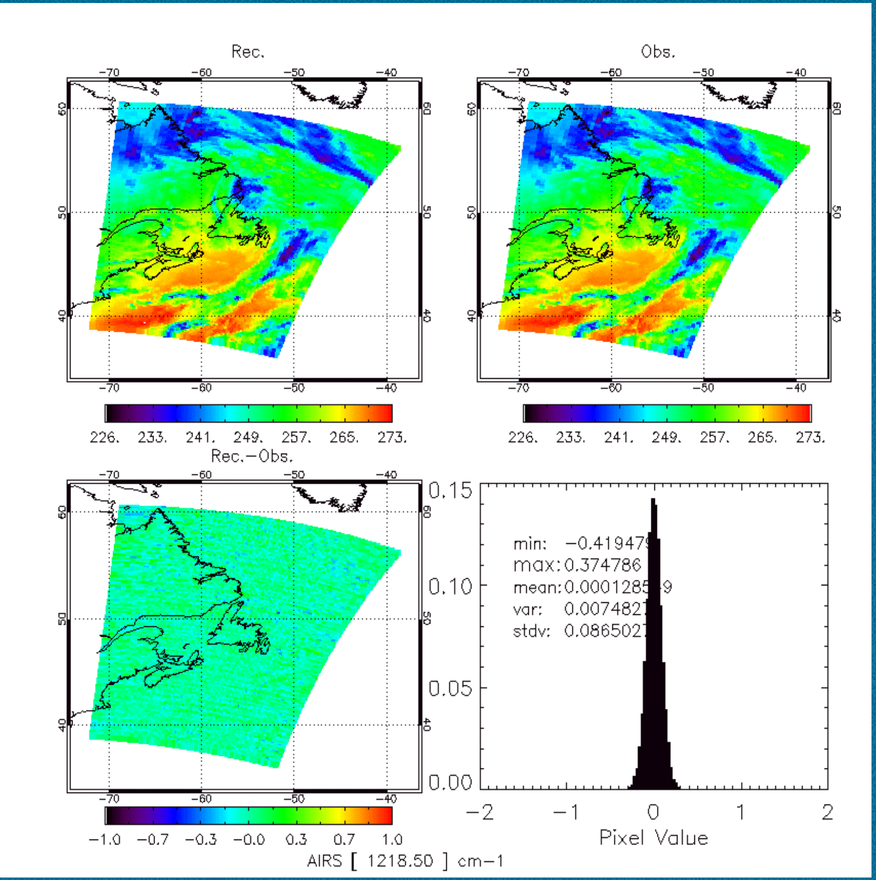
The difference between reconstructed minus observed is not too dependent on eigenvector set, but granule eigenvectors will improve the compression ratio by at least 2x and will also better account for 3 sigma events (volcanic eruptions)



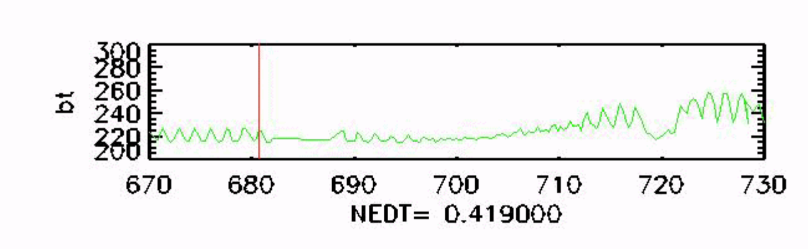


CH4 and H2O

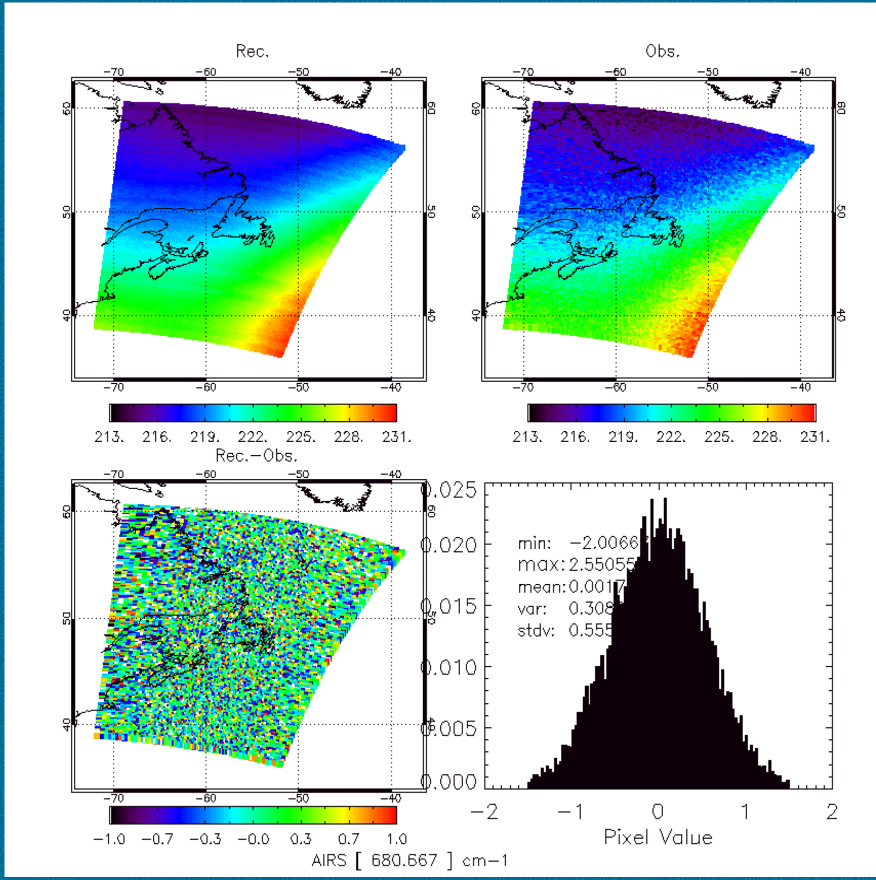

bt
300
280
260
240
220
1200
1220
1240
1260
1280
1300
NEDT= 0.108000


Rec.
Obs.
226. 233. 241. 249. 257. 265. 273.
226. 233. 241. 249. 257. 265. 273.
Rec.−Obs.
−1.0 −0.7 −0.3 −0.0 0.3 0.7 1.0
min: −0.41947
max: 0.374786
mean: 0.0001285
var: 0.007482
stdv: 0.086502
0.15
0.10
0.05
0.00
−2 −1 0 1 2
Pixel Value
AIRS [1218.50] cm−1

Q-Branch

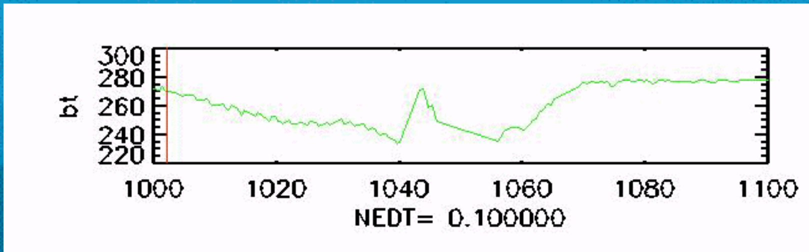


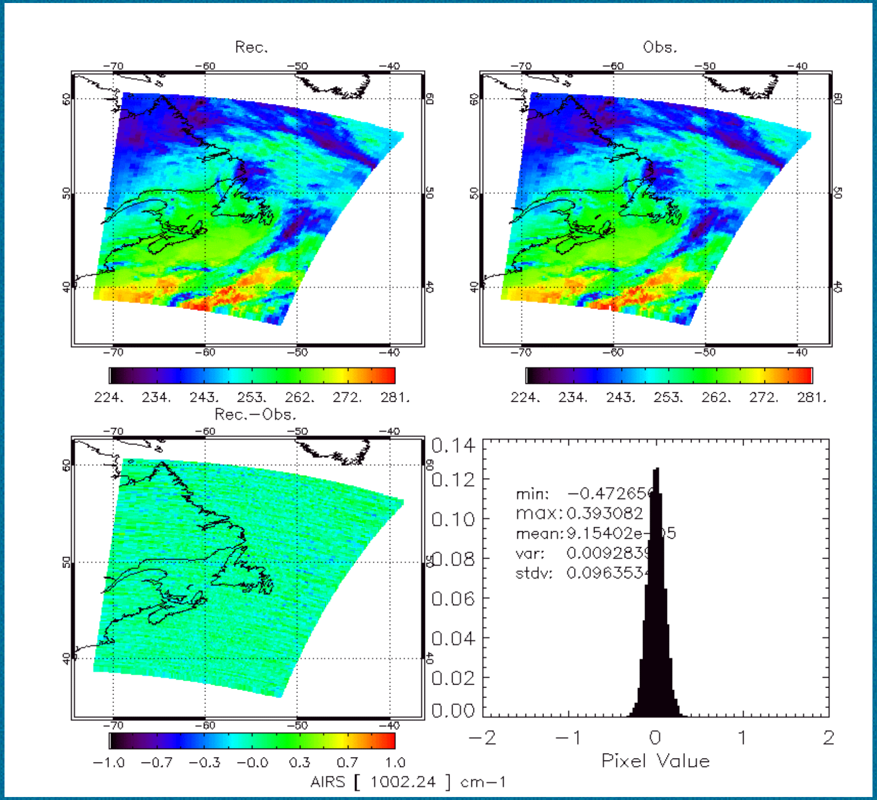
bt
300
280
260
240
220
200
670
680
690
700
710
720
730
NEDT= 0.419000



Rec.
Obs.
213. 216. 219. 222. 225. 228. 231.
213. 216. 219. 222. 225. 228. 231.
Rec.−Obs.
−1.0 −0.7 −0.3 −0.0 0.3 0.7 1.0
AIRS [680.667] cm−1
0.025
0.020
0.015
0.010
0.005
0.000
−2 −1 0 1 2
Pixel Value

Ozone Band





PCA compression seems to be working quite well

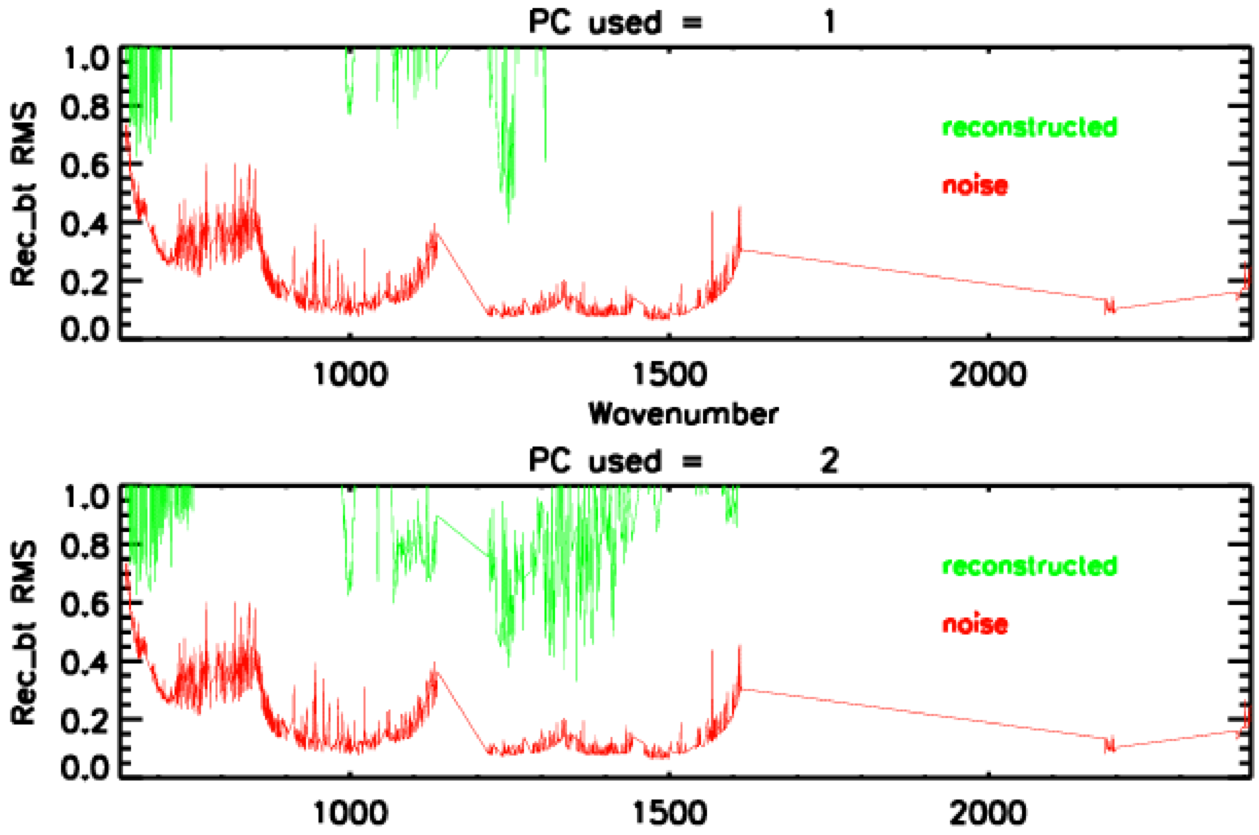
The residuals are at noise levels and can be compressed and stored in a separate file for lossless compression

Most people will not want the residuals.

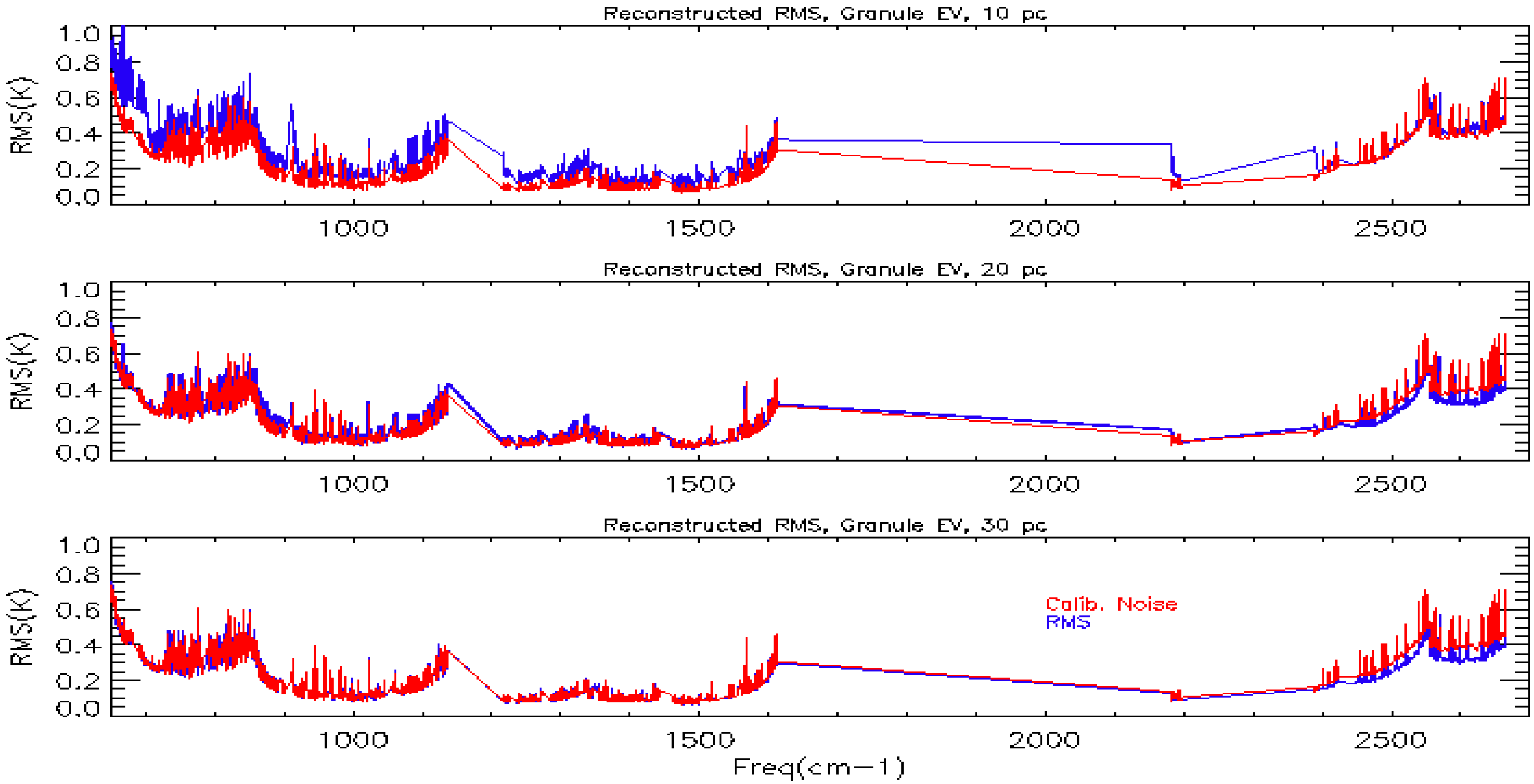
The picture to the left can be also used as a form of metadata which will convince the user that the lossy compression is favorable

Users can decide whether they want the residual file

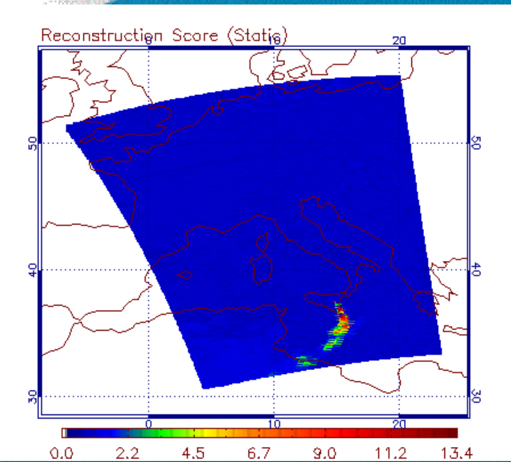
Reconstruction RMS vs. # of PC

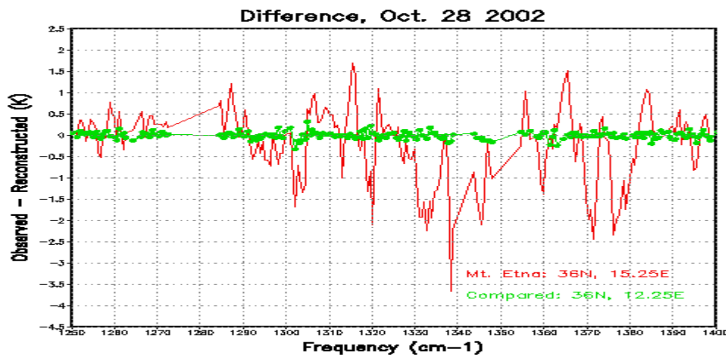


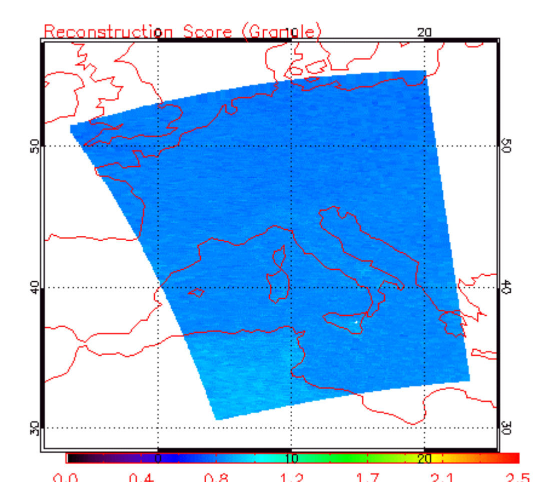
Reconstructed error vs. the AIRS Instrument Noise

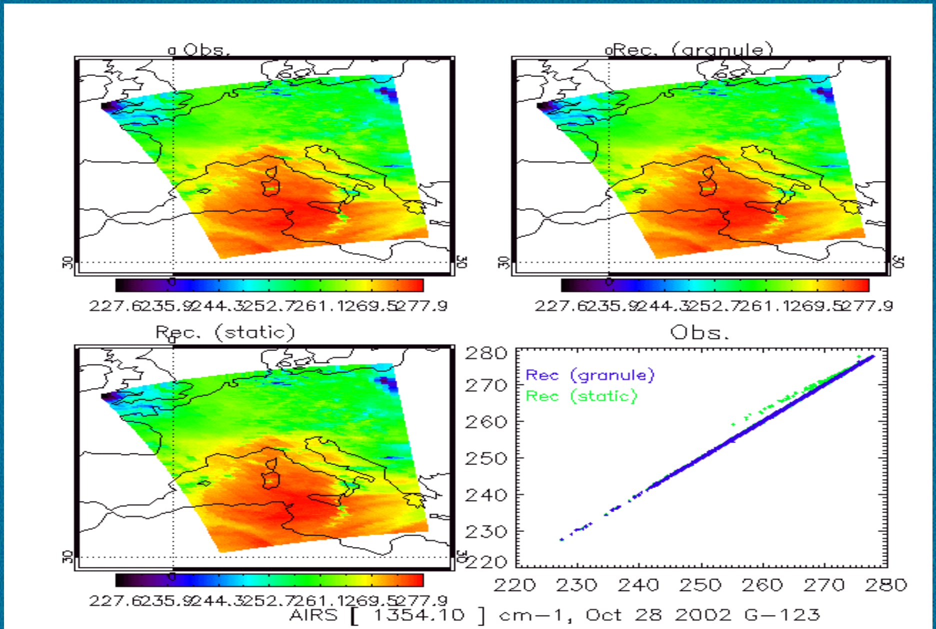


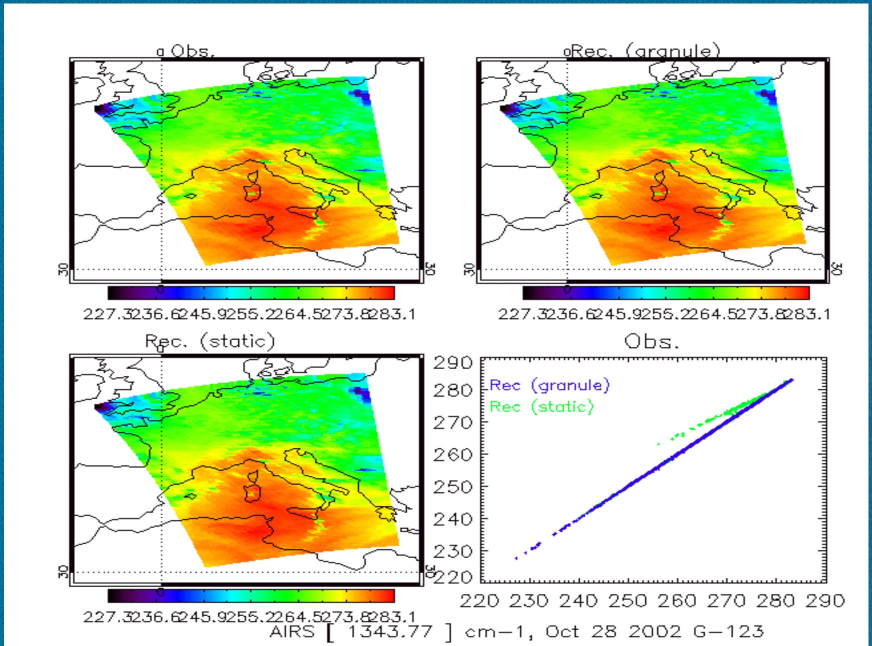
Mt. Etna Volcano Eruption Episode











Granule EOFs – Pros and Cons

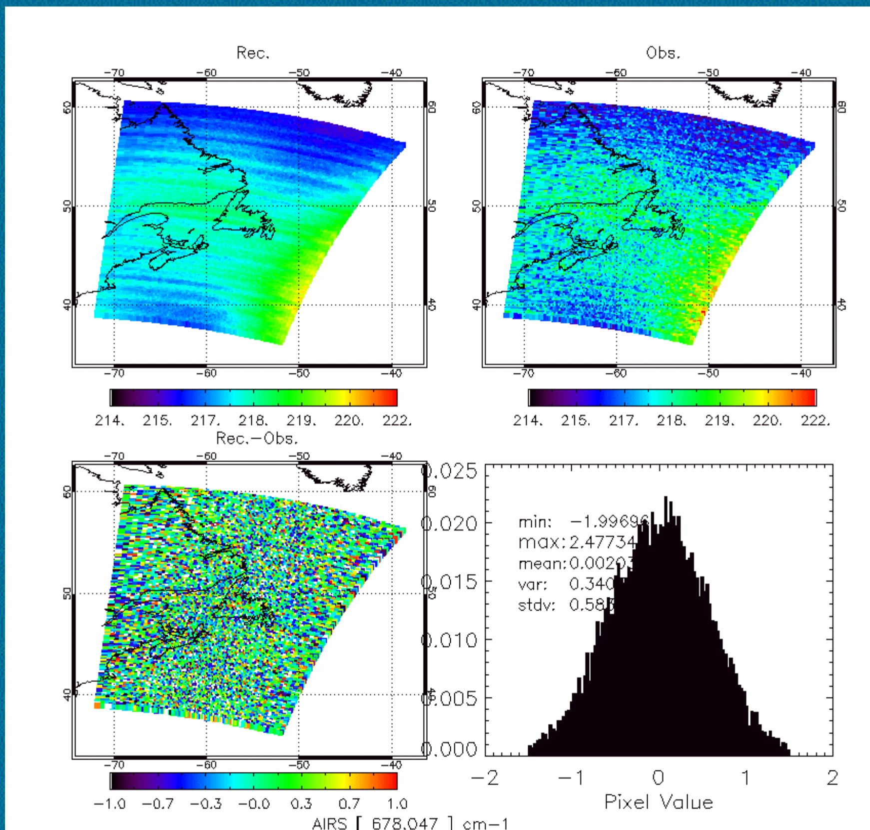
Granule eofs will provide better compression ratios ~ 50-100x

Granule eofs are more time consuming. (45 minutes per granule)

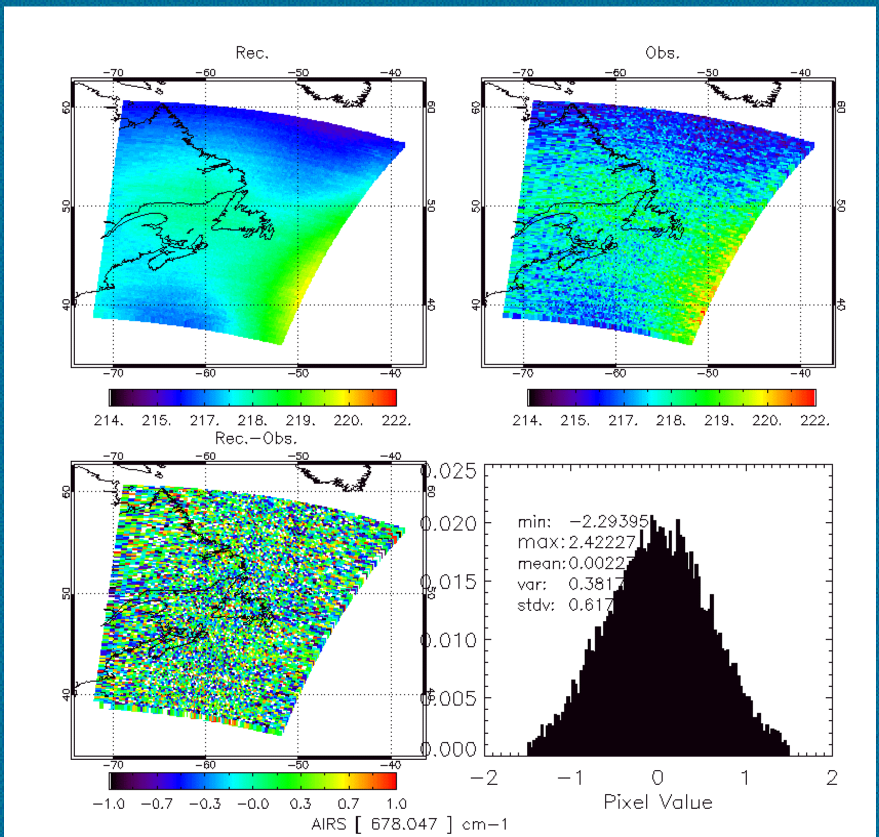
Granule eofs may reproduce striping.

Too many EOFs will reproduce striping

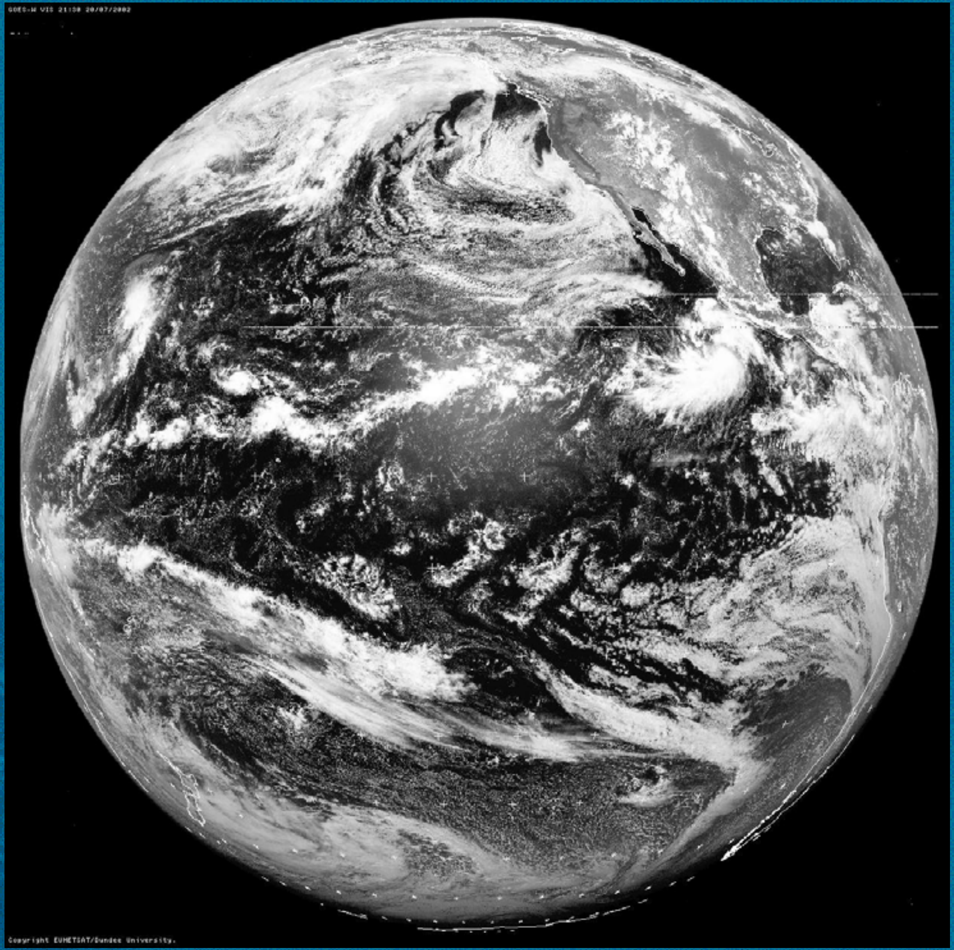
80 pc



40 pc



Cloud-Cleared Radiances

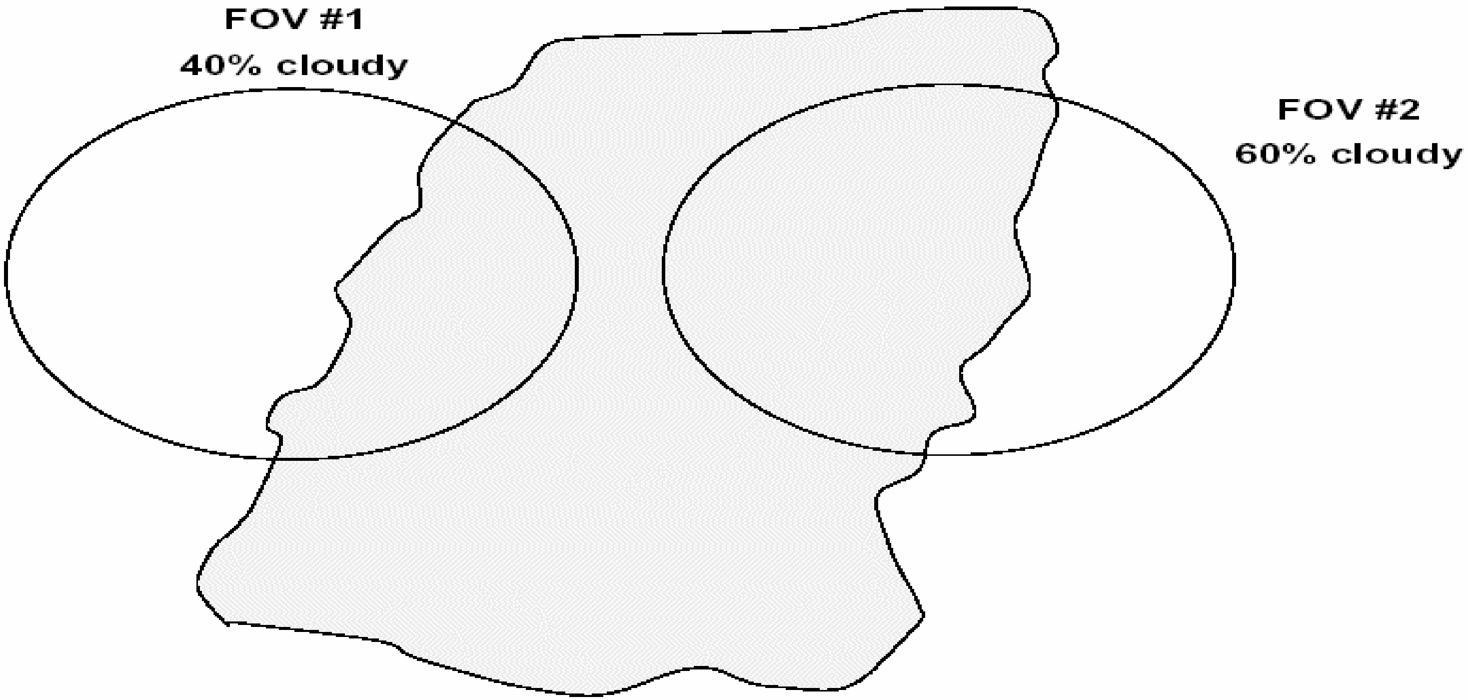


Cloud Clearing

$$R1 = (1-a1)*R1clear + a1*Rcloud$$

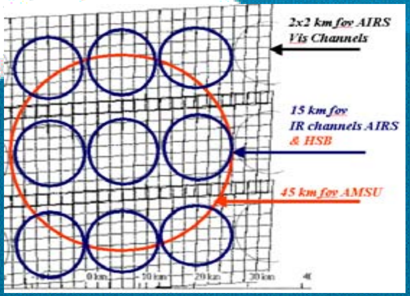
$$R2 = (1-a2)*R2clear + a2*Rcloud$$

Assume Scene Is Identical in FOV's except Fraction of Cloud



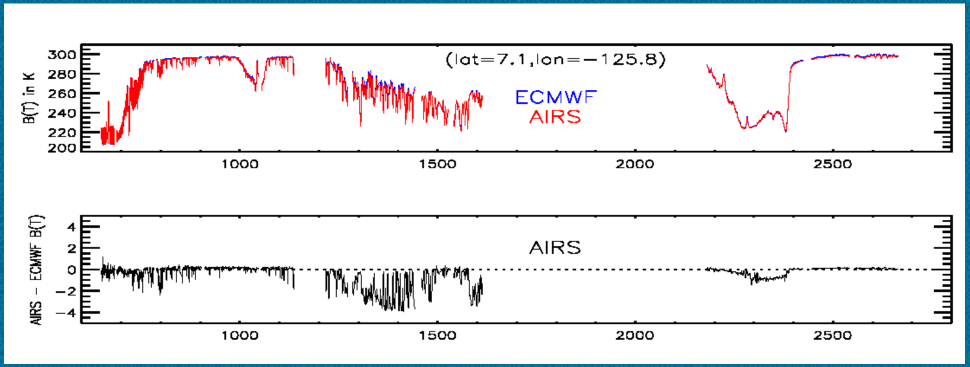
Two AIRS field of views (FOV's) are illustrated showing that each FOV has some fraction of clear radiance and some fraction of cloudy radiance. We define the ensemble of FOV's as the retrieval field of regard (FOR).

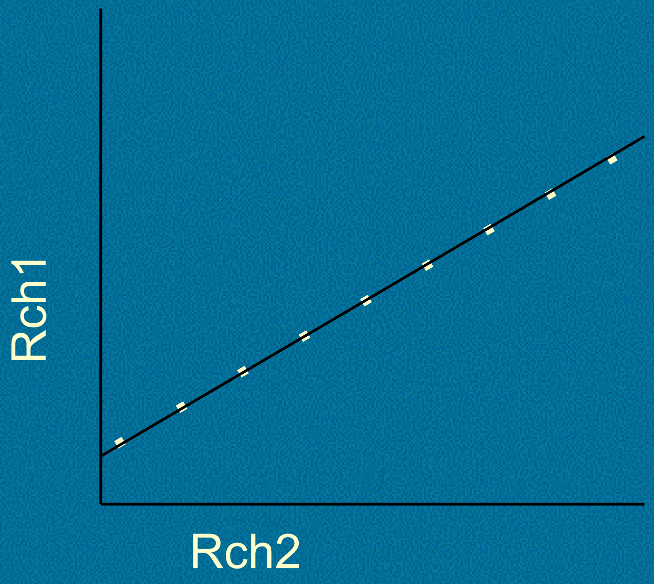
$$Rclear(i) = R1(i) + N * [R1(i)-R2(i)] \quad i = \text{channel \#}$$



$$N = a1/(a2-a1)$$

$$N = (Rclear - R1)/(R1 - R2)$$



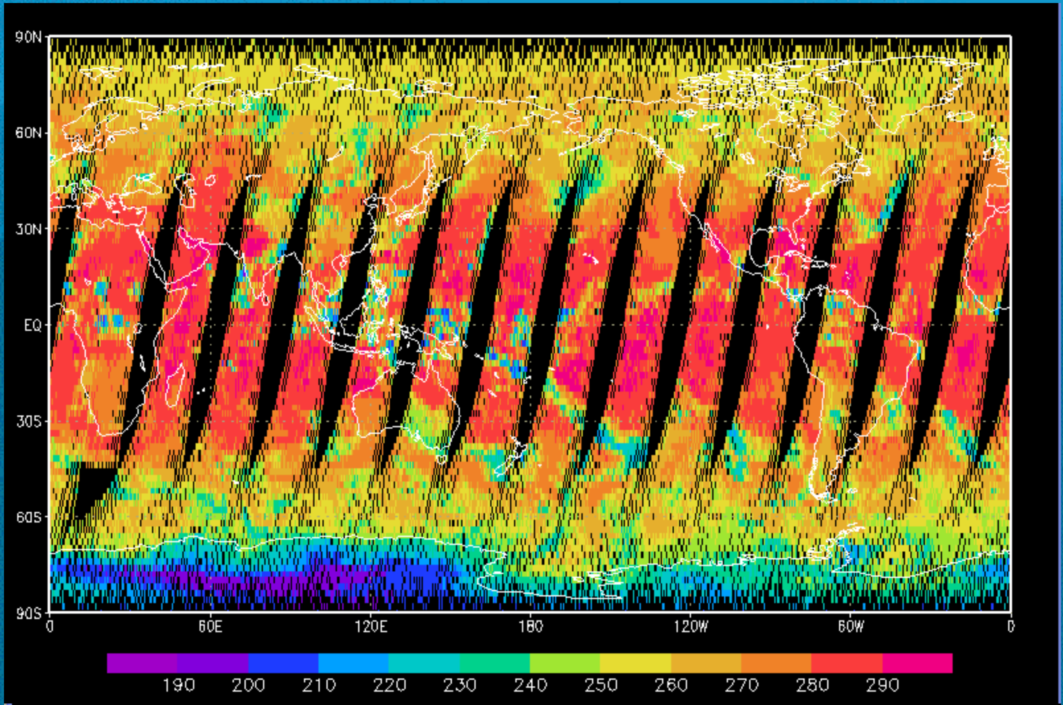


$$Rclear(i) - Rmean(i) + \sum N(j)*[R(i,fov\#)-Rmean(i)]$$

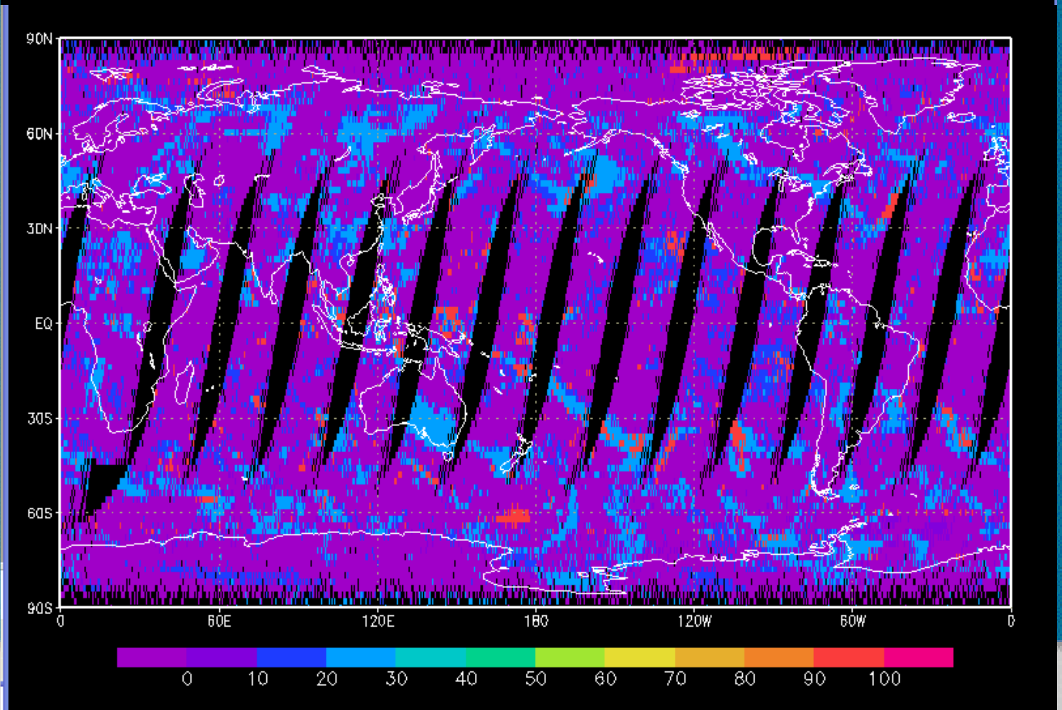
j = cloud formation

Raw

1000 cm-1



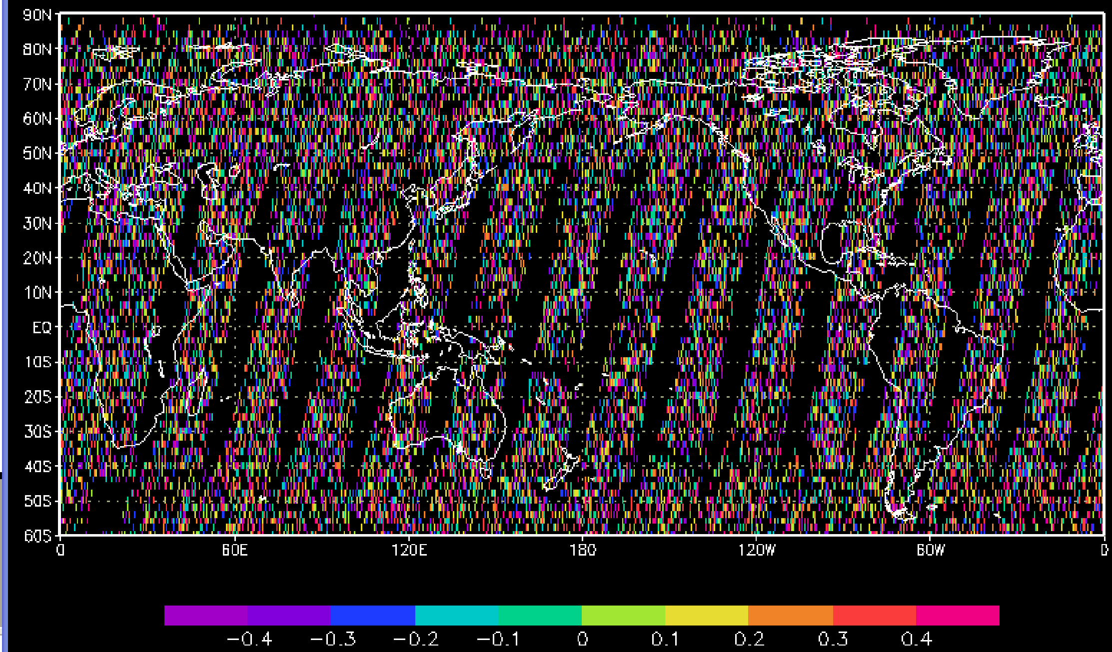
Cloud-Cleared



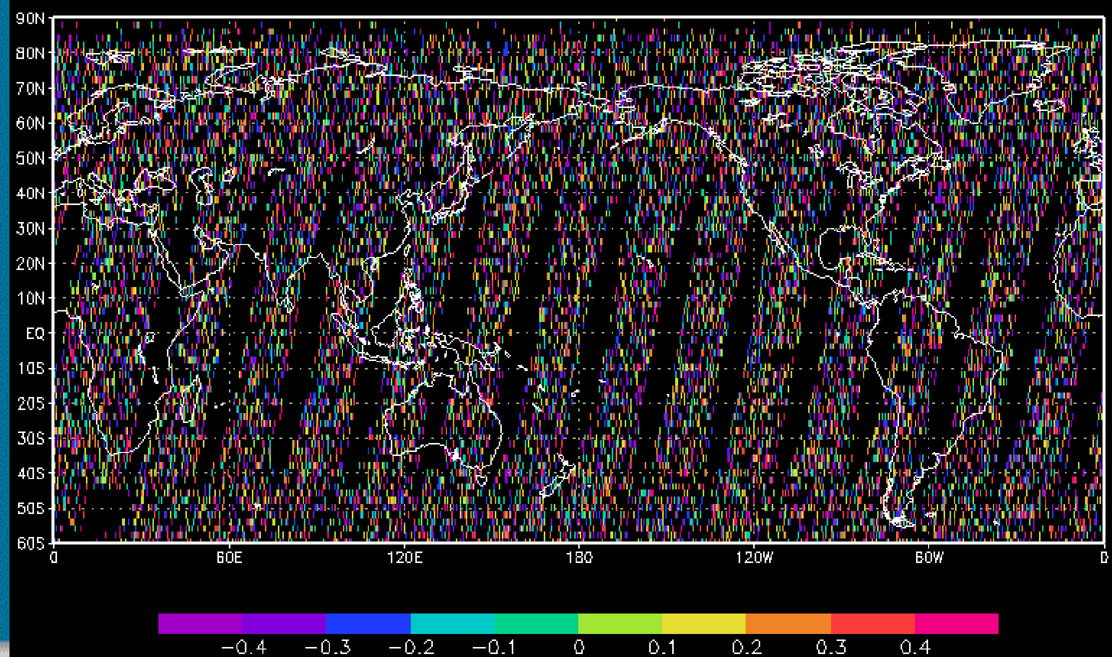
ALL diff < +- 0.5 K

649.61 cm-1 (peak ~ 70 mb)

Cloud-cleared minus ECMWF



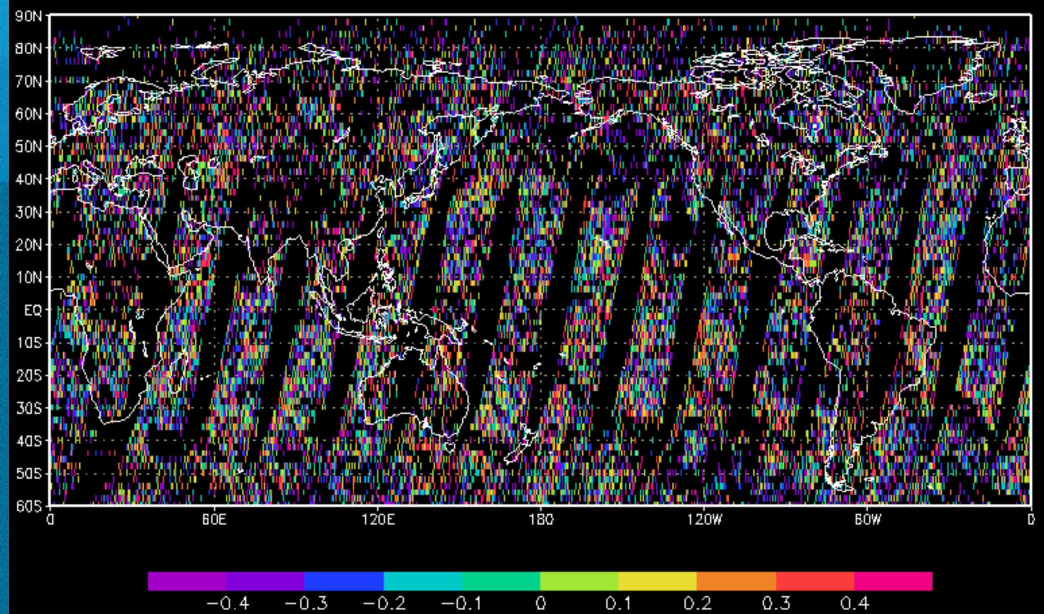
Raw minus ECMWF

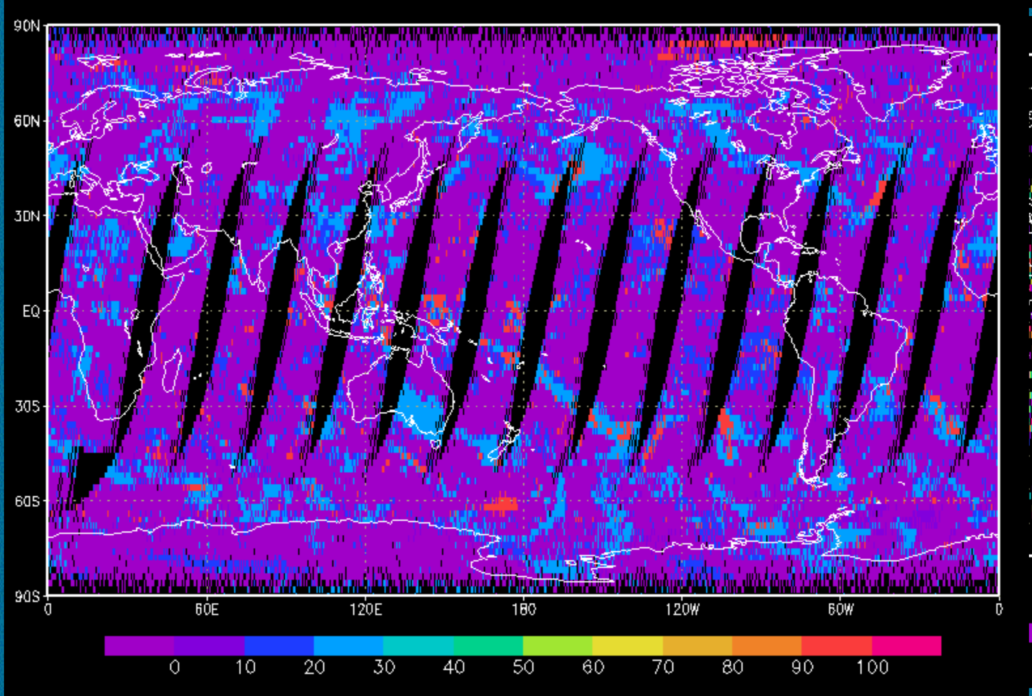


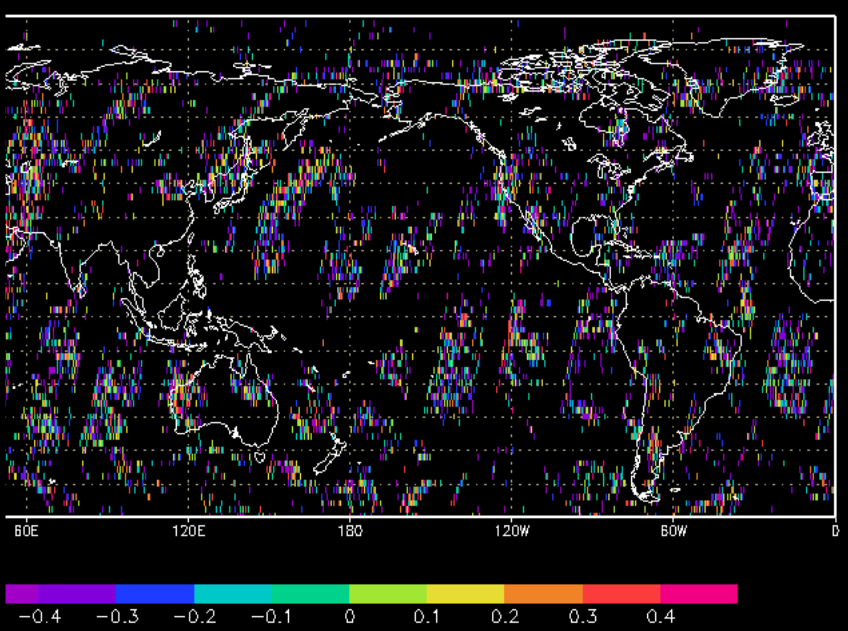
ALL diff < +- 0.5 K

735.69 cm-1 (peak ~ 700 mb)

Cloud-cleared minus ECMWF



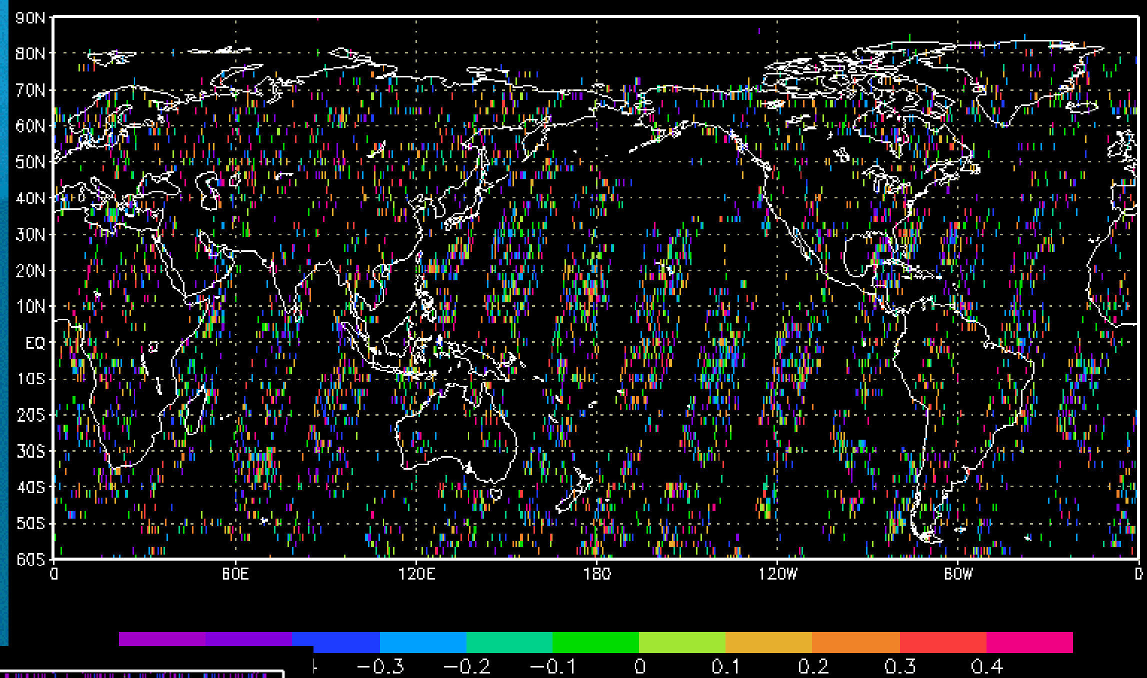


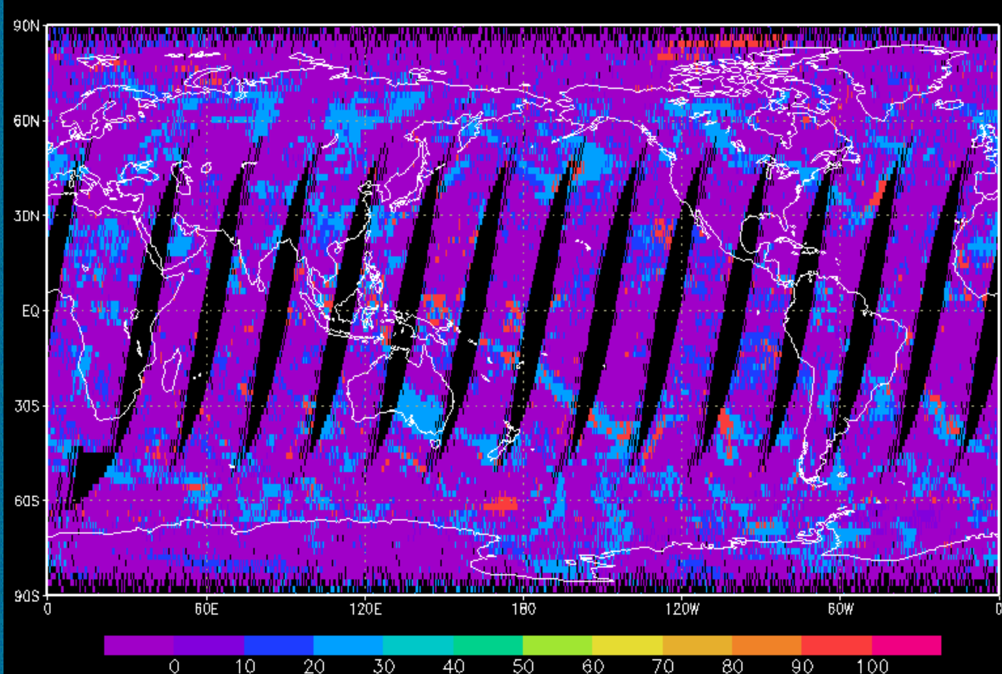


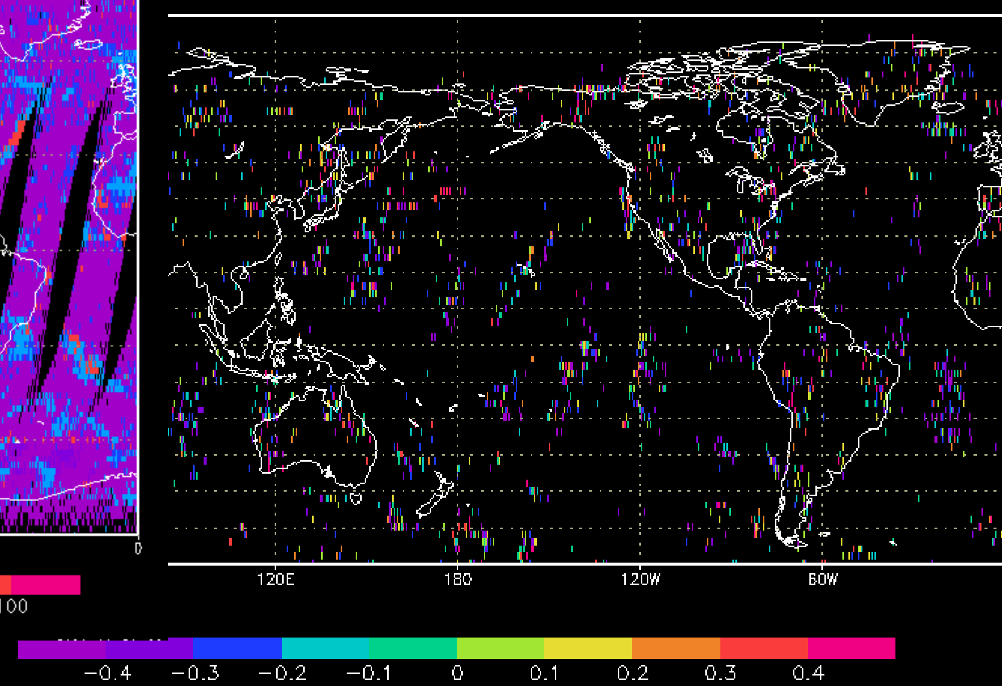
ALL diff < +- 0.5 K

965 cm-1   (window)

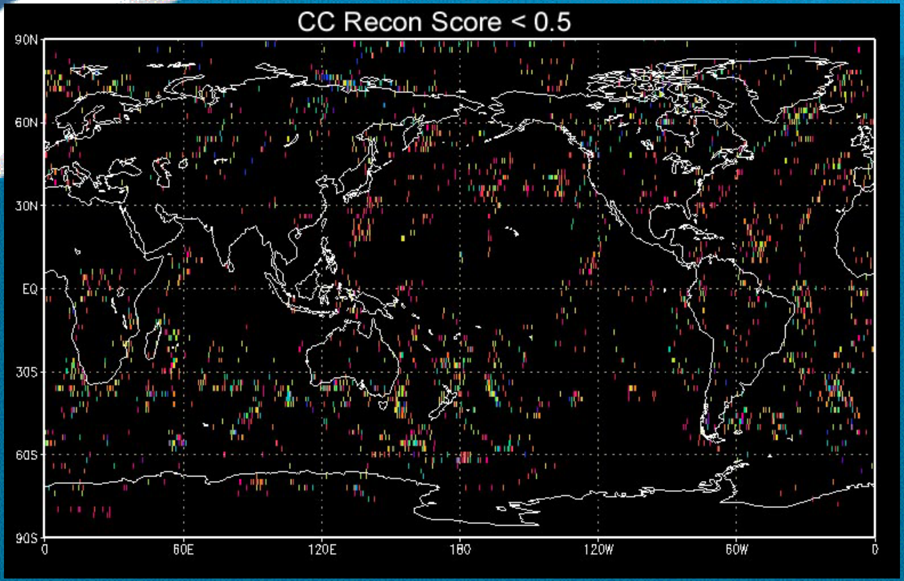
Cloud-cleared minus ECMWF

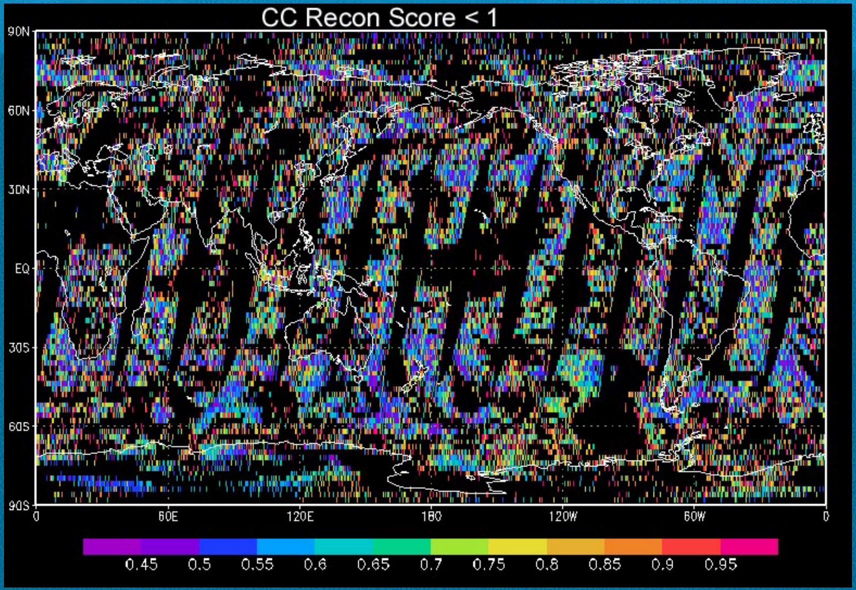


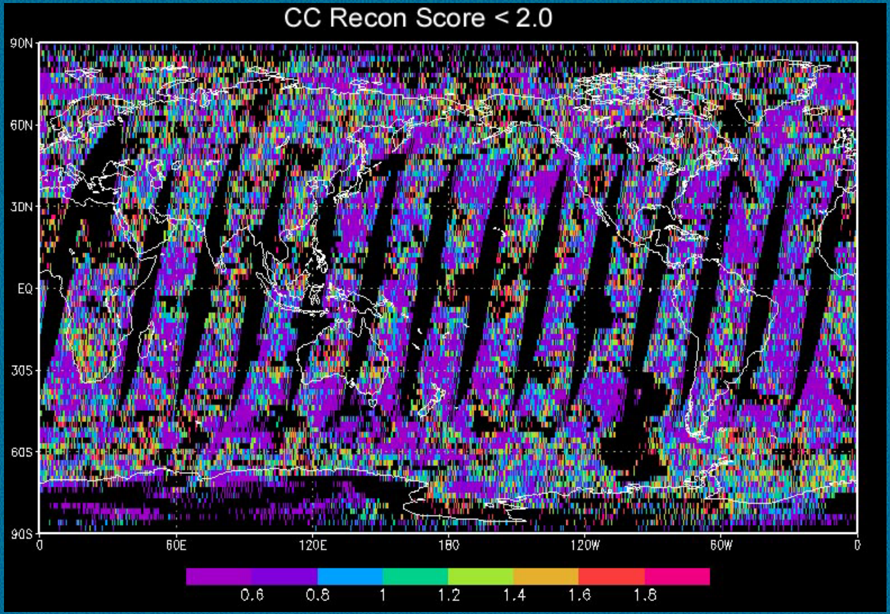


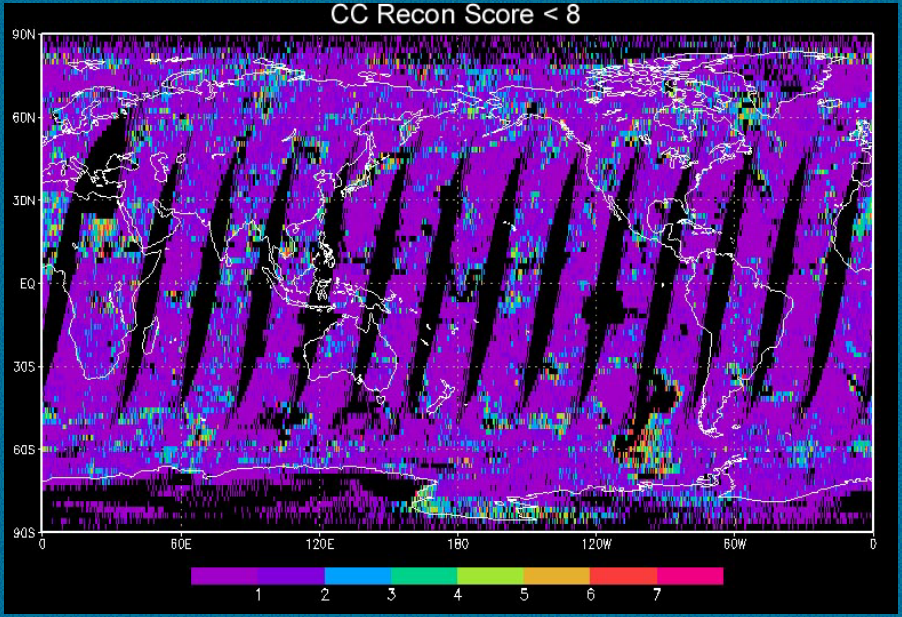


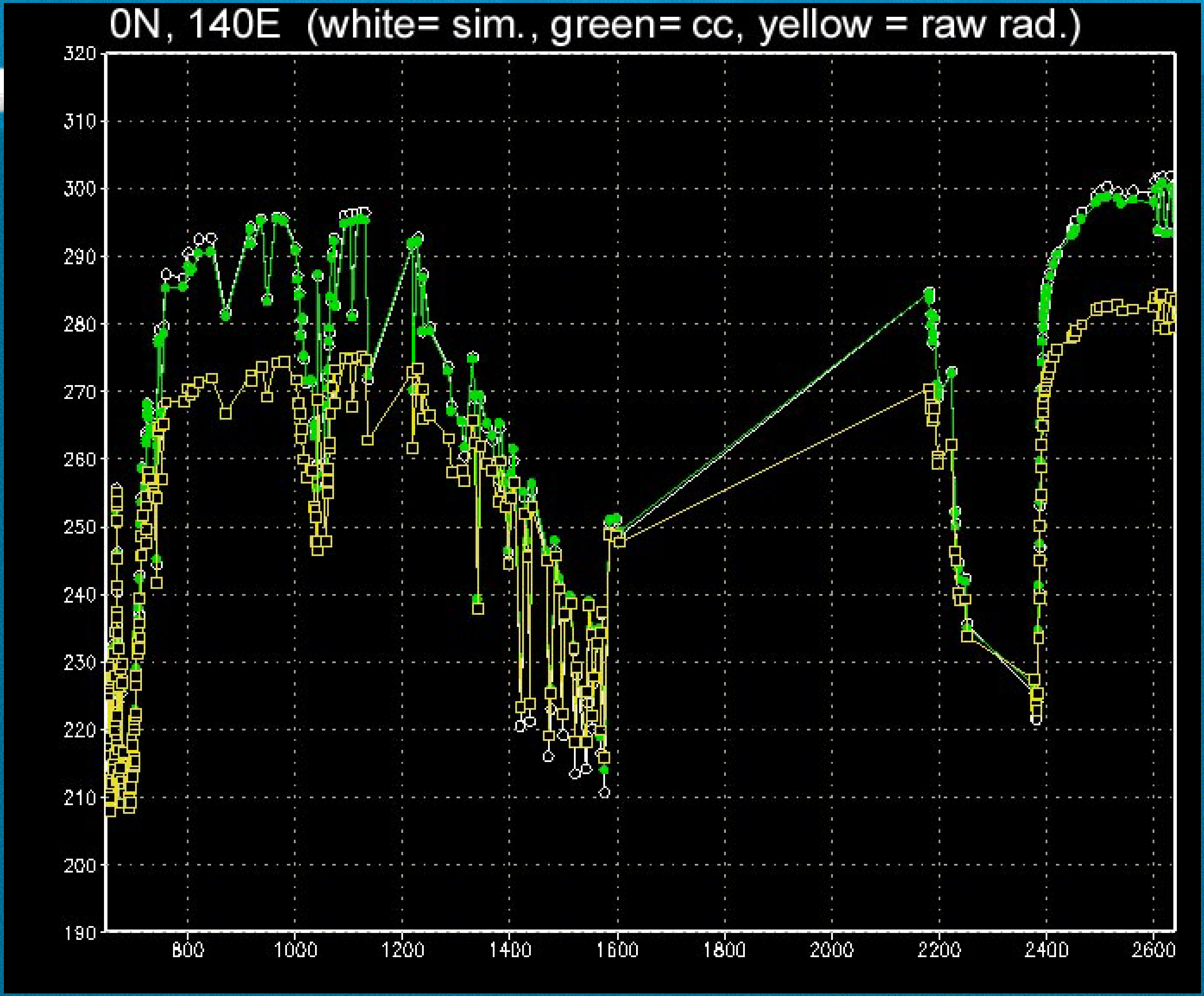
CC NOISE AMPLIFICATION MAP

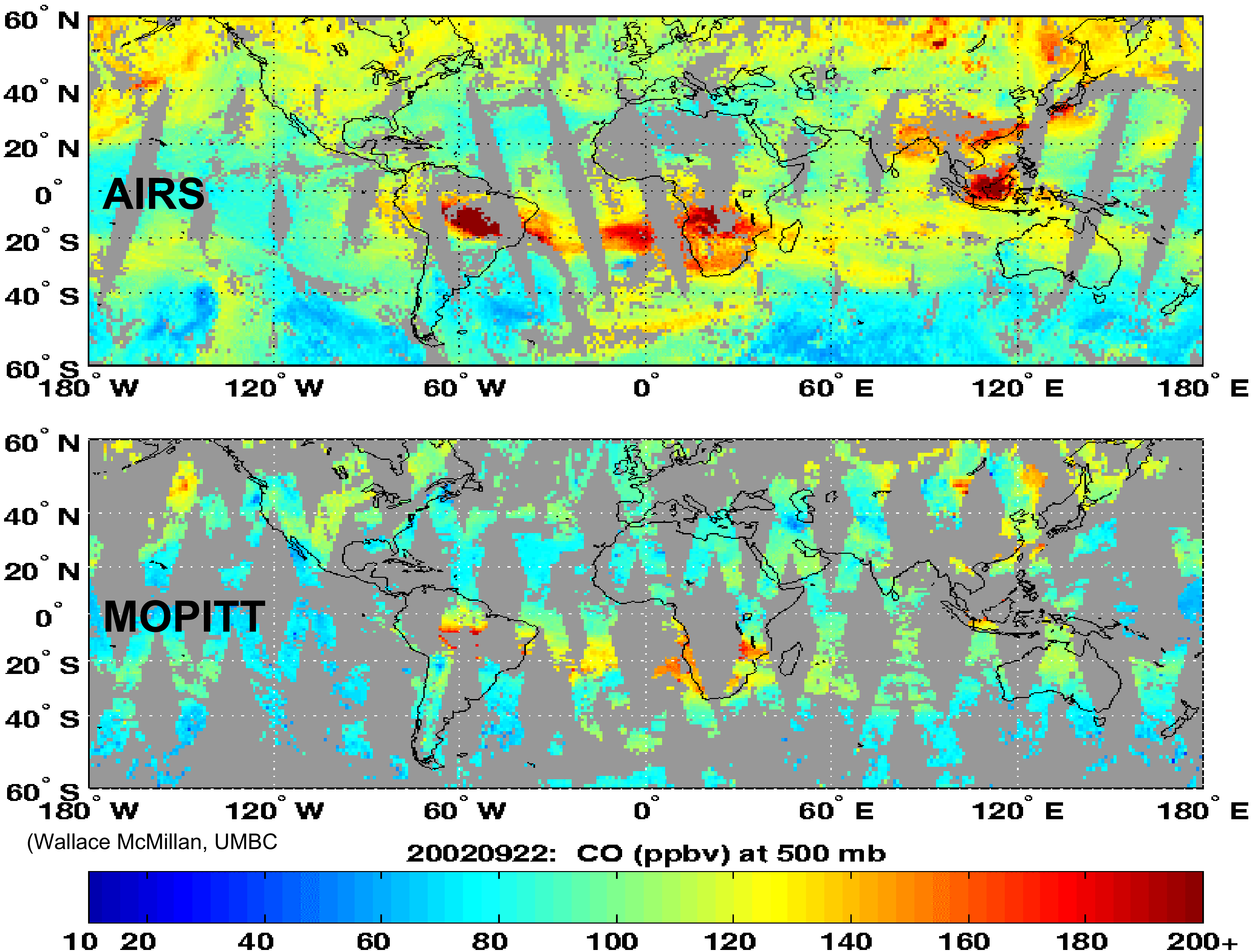








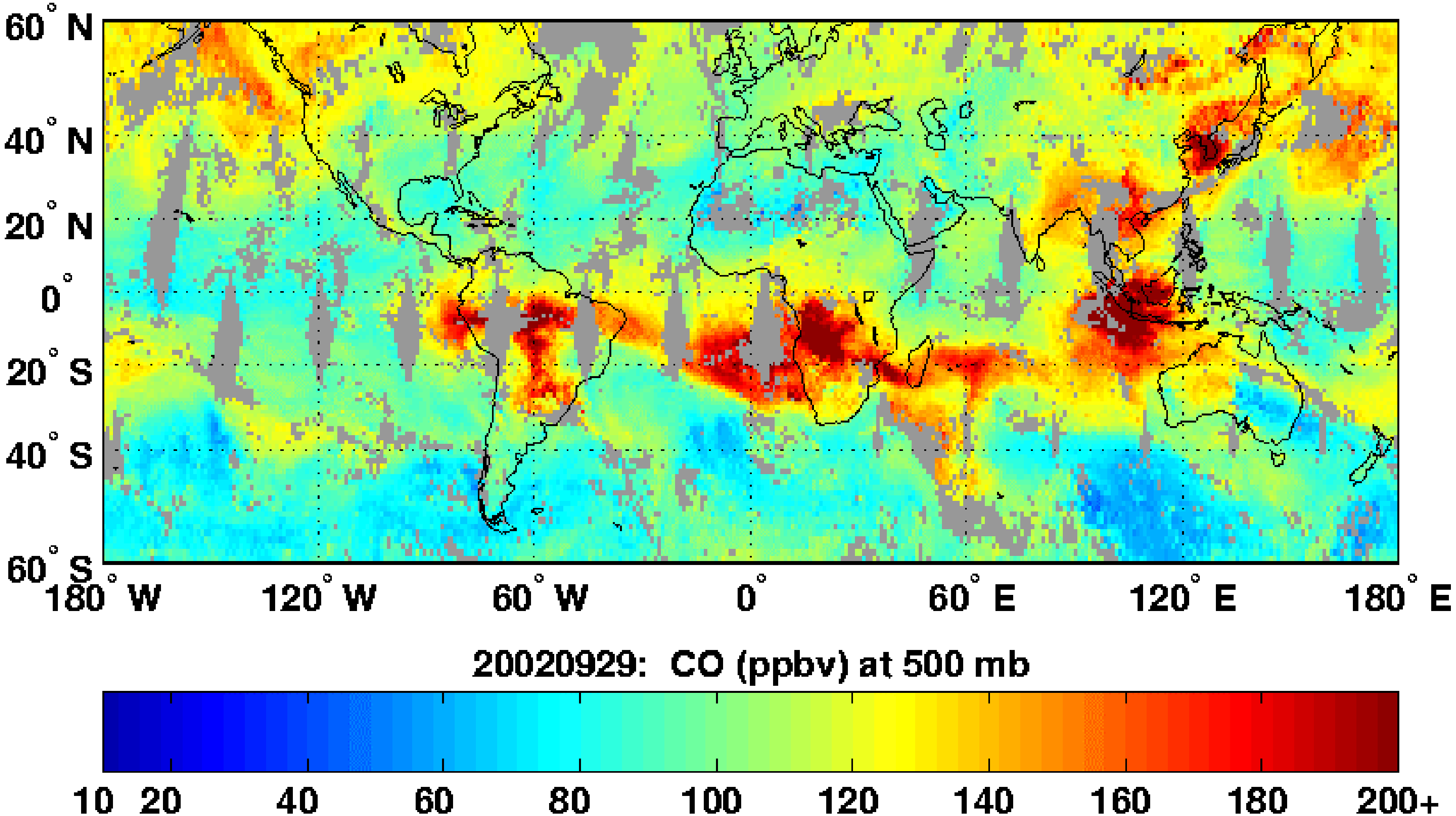

0N, 140E (white= sim., green= cc, yellow = raw rad.)
320
310
300
290
280
270
260
250
240
230
220
210
200
190
800
1000
1200
1400
1600
1800
2000
2200
2400
2600

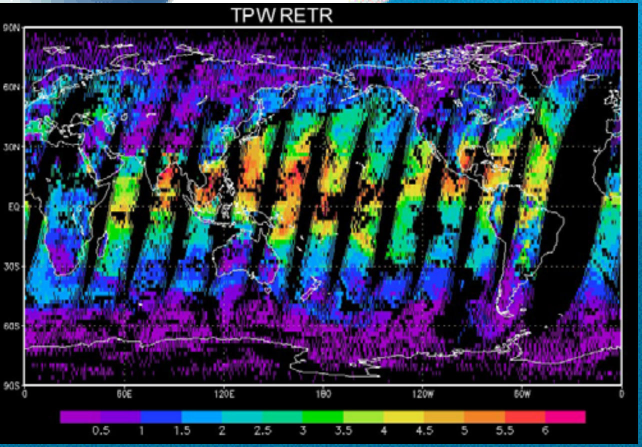


AIRS
MOPITT
60° N
40° N
20° N
0°
20° S
40° S
60° S
180° W
120° W
60° W
0°
60° E
120° E
180° E
(Wallace McMillan, UMBC
20020922: CO (ppbv) at 500 mb
10 20 40 60 80 100 120 140 160 180 200+

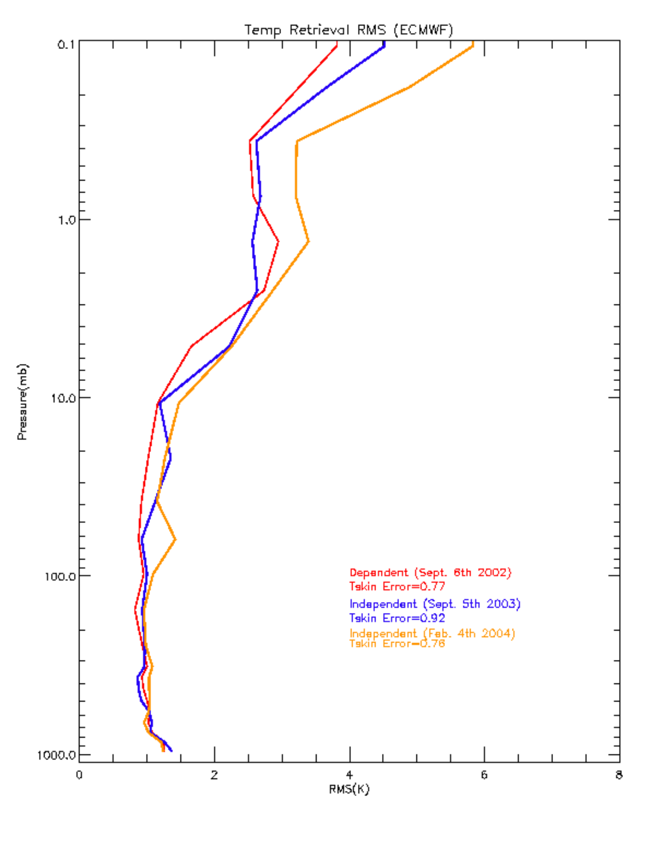
AIRS Global CO Movie

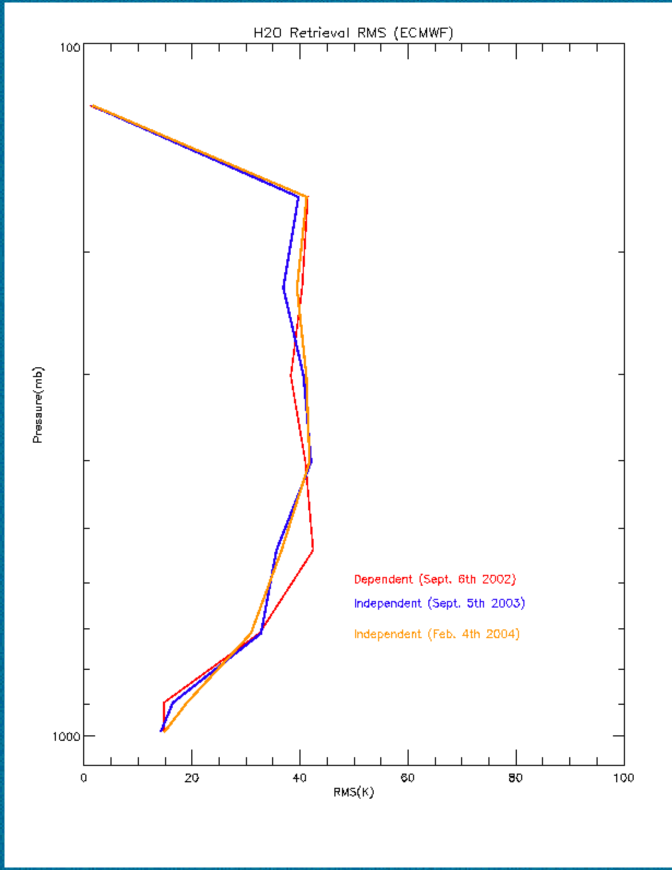
(Wallace McMillan, UMBC)





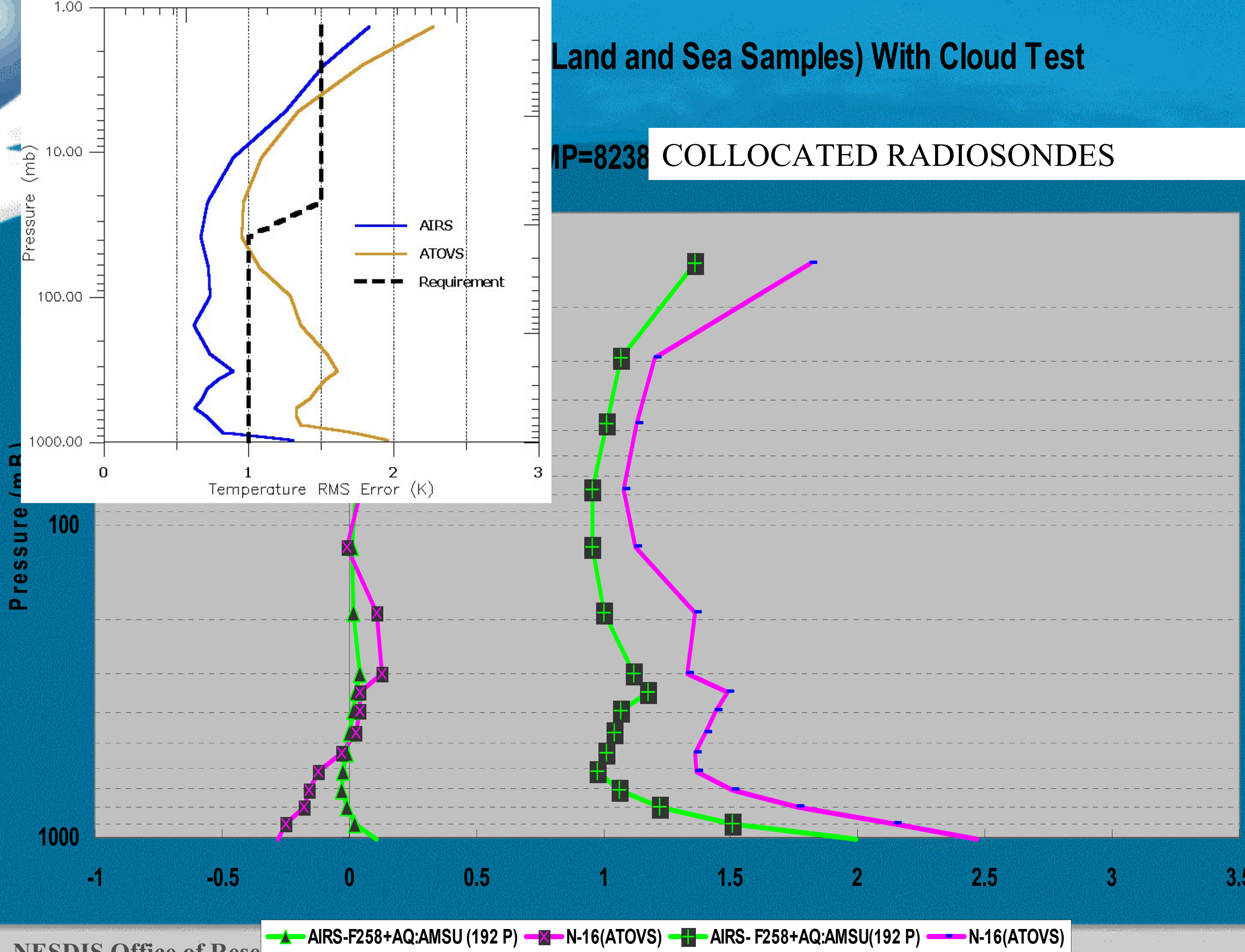
RMS Differences (retrieval minus ECMWF)





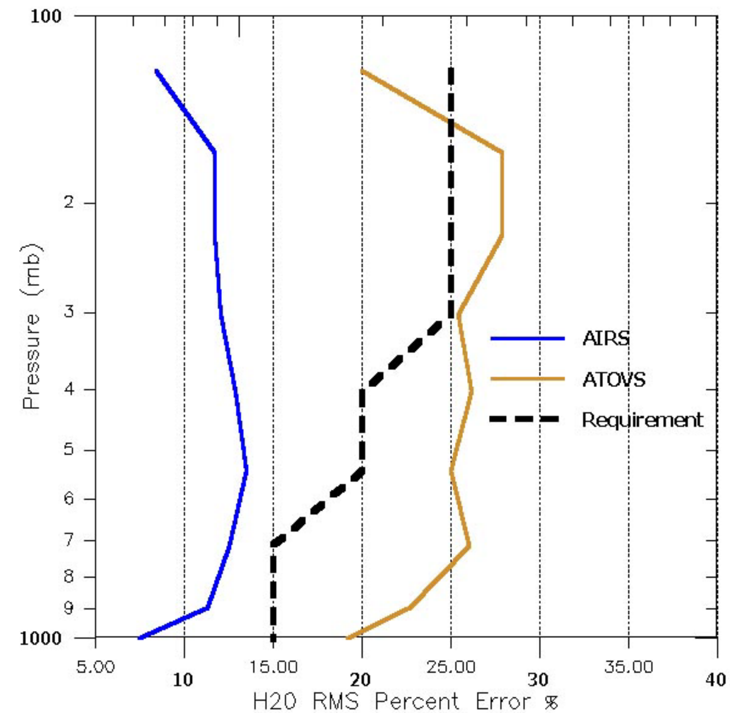
Land and Sea Samples) With Cloud Test

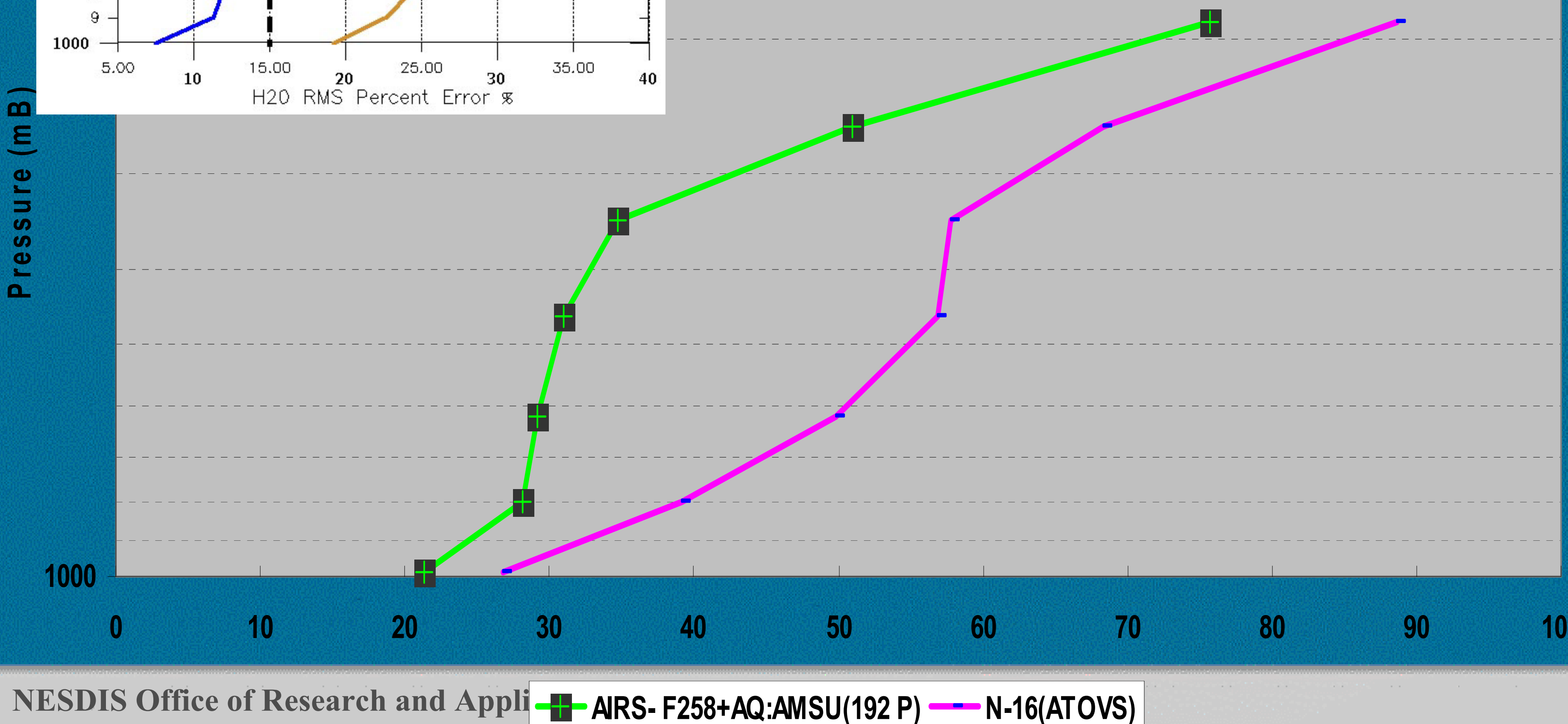
MP=8238 COLLOCATED RADIOSONDES



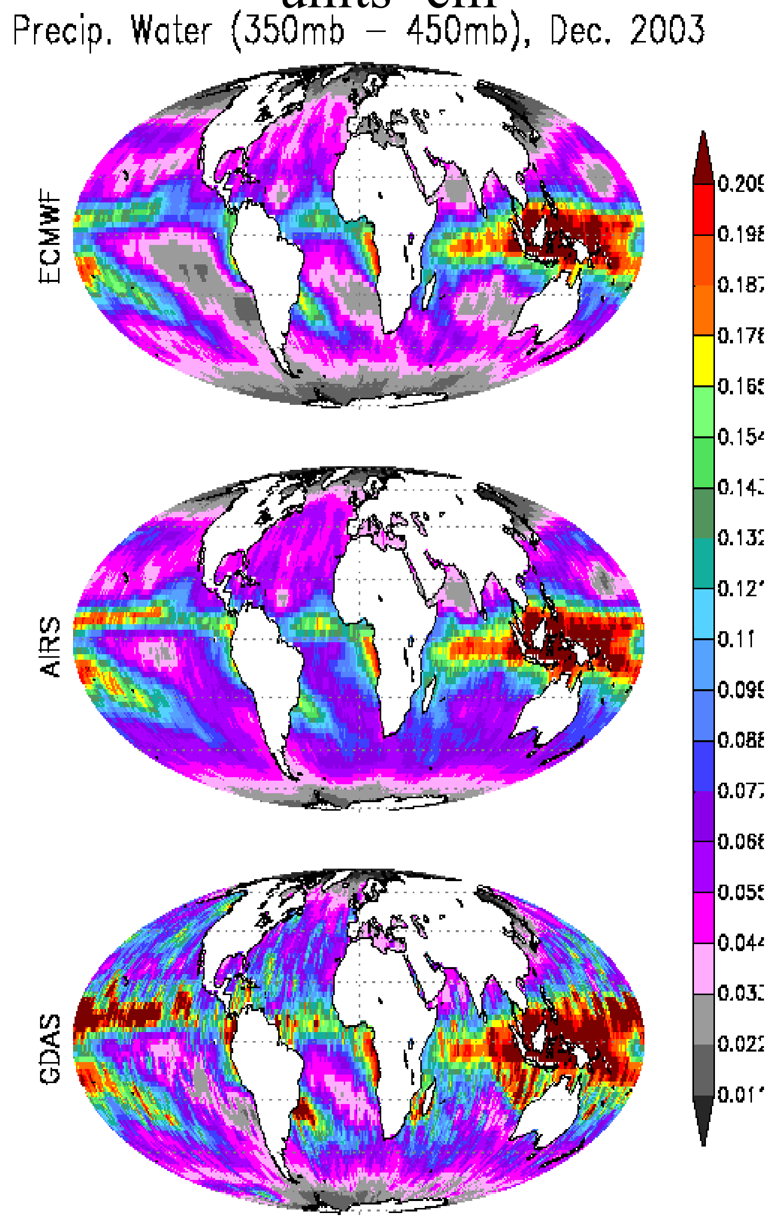
d and Sea Samples) With Cloud Test

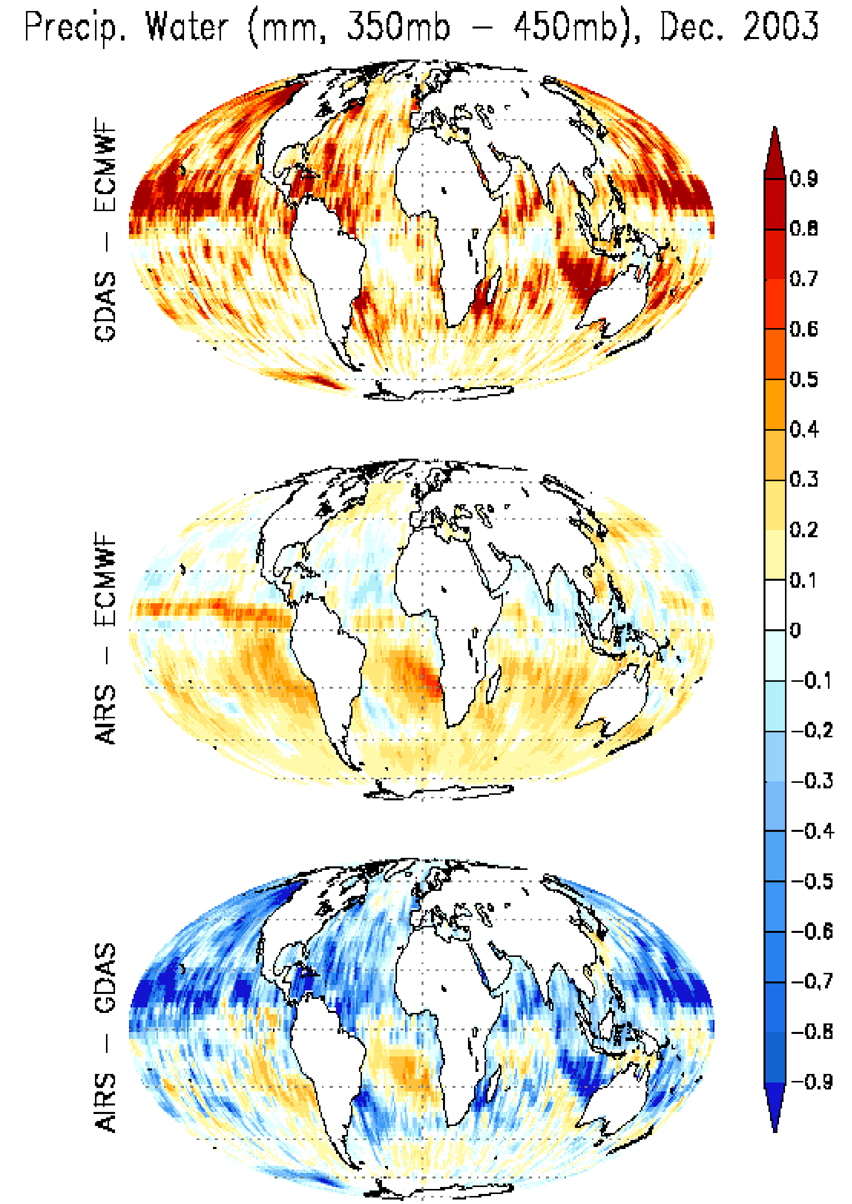
238 COLLOCATED RADIOSONDES

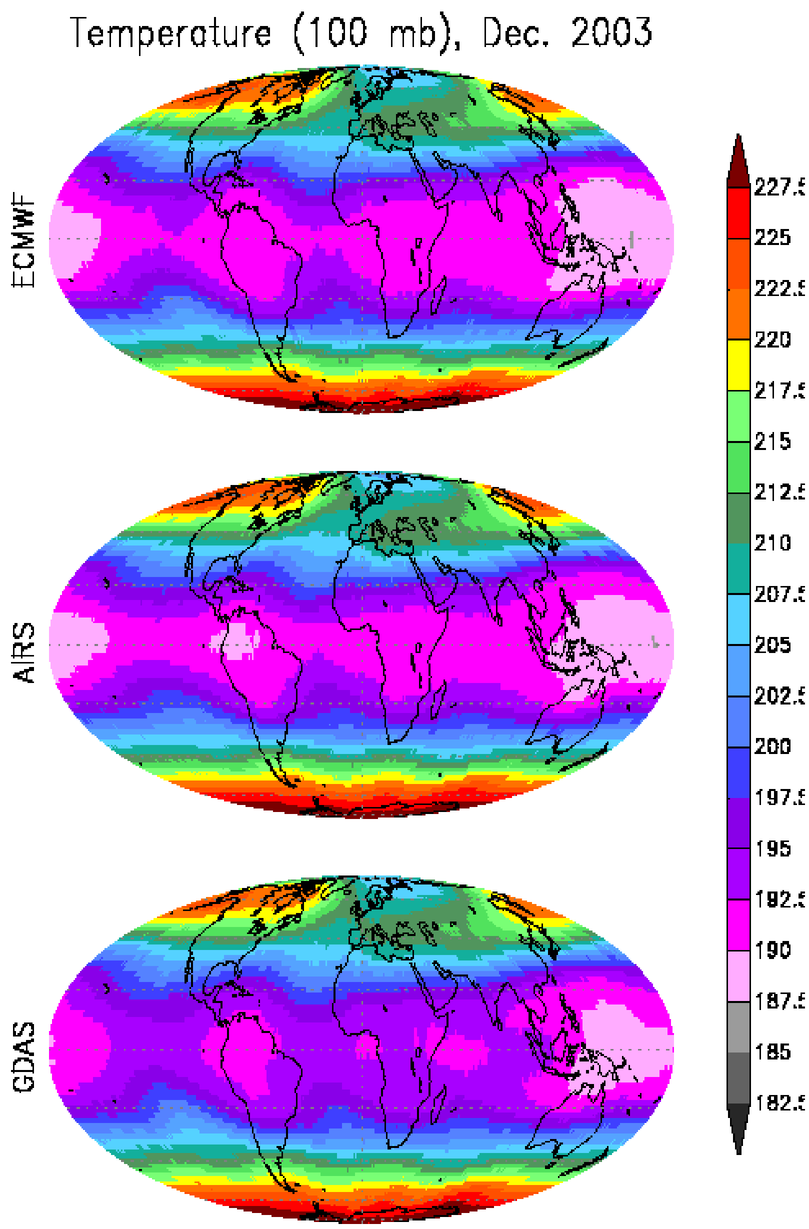


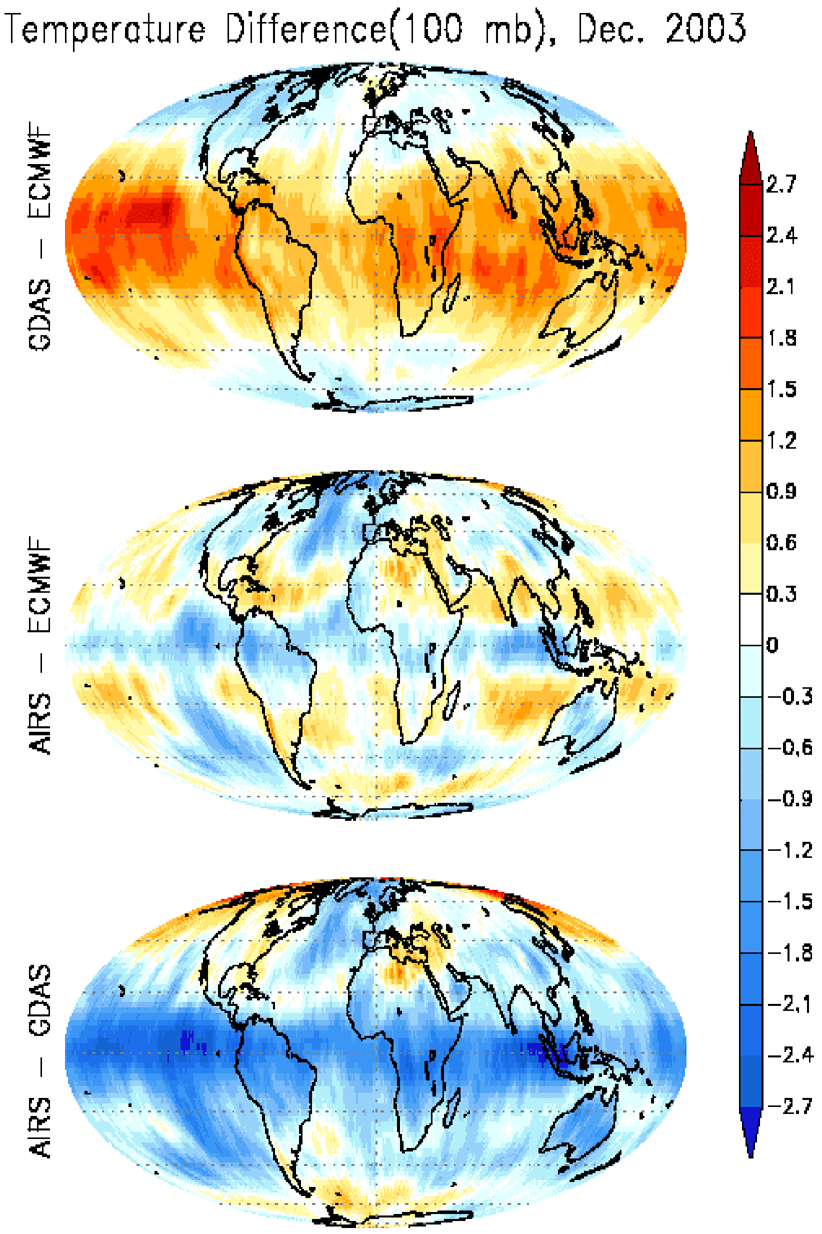


units=cm


Precip. Water (350mb – 450mb), Dec. 2003
ECMWF
AIRS
GDAS
0.11


Precip. Water (mm, 350mb – 450mb), Dec. 2003
GDAS – ECMWF
AIRS – ECMWF
AIRS – GDAS
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
–0.1
–0.2
–0.3
–0.4
–0.5
–0.6
–0.7
–0.8
–0.9


Temperature (100 mb), Dec. 2003
ECMWF
AIRS
GDAS
227.5
225
222.5
220
217.5
215
212.5
210
207.5
205
202.5
200
197.5
195
192.5
190
187.5
185
182.5


Temperature Difference(100 mb), Dec. 2003
GDAS – ECMWF
AIRS – ECMWF
AIRS – GDAS
2.7
2.4
2.1
1.8
1.5
1.2
0.9
0.6
0.3
0
-0.3
-0.6
-0.9
-1.2
-1.5
-1.8
-2.1
-2.4
-2.7

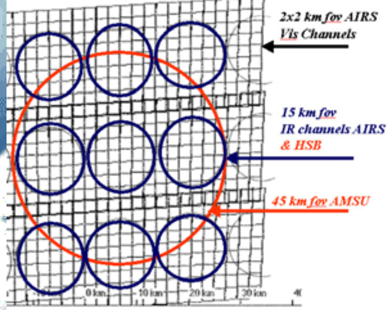
AIRS Assimilation -The Next Steps

(Including AMSU/MODIS/MFG….)

* All data plus data selection / thinning studies

** all channels plus channel selection / noise red. studies

Data utilised (AQUA)	Spatial Res.	Spectral Res.	Comment
AIRS	Full* all data plus data selection / thinning studies	Current 300 Ch.	Current 3DVar CLR Rd assim
AIRS	Full*	Current 300 Ch. Recon.Rads	Current 3DVar CLR Rd assim
AIRS AMSU and MODIS	Full*	300 Cld Cleared Rads.	AMSU/MODIS used in QC
AIRS AMSU and MODIS	Full*	Full** all channels plus channel selection / noise red. studies	Current 3DVar CLR Rd assim
AIRS AMSU MODIS	Full*	Full**	Cloudy Rads Used



AIRS Radiance Products

Spectrally and Spatially Thinned Radiances

- Center of every AMSU fov
- 3 x 3 array from every other AMSU fov

Principal Component Scores (Spatially Thinned)

Reconstructed radiances (Spatially/Spectrally Thinned)

Cloud-Cleared radiances (Spatially/Spectrally Thinned)

Cloud-Cleared PCS (Spatially Thinned)

Clear MODIS @ AIRS resolution

Summary - Roadmap to larger impacts

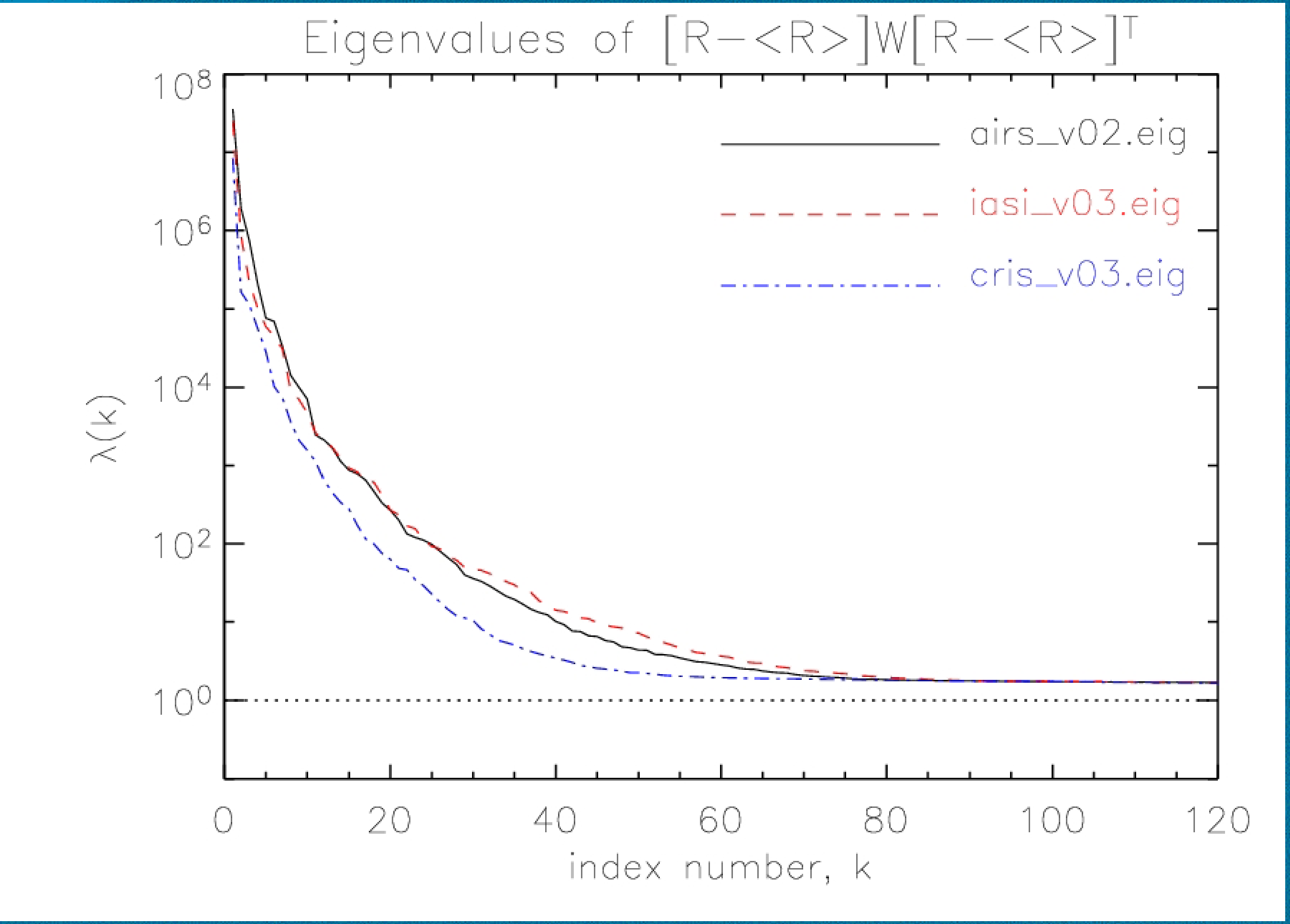
Use reconstructed radiances

Use cloud cleared radiances

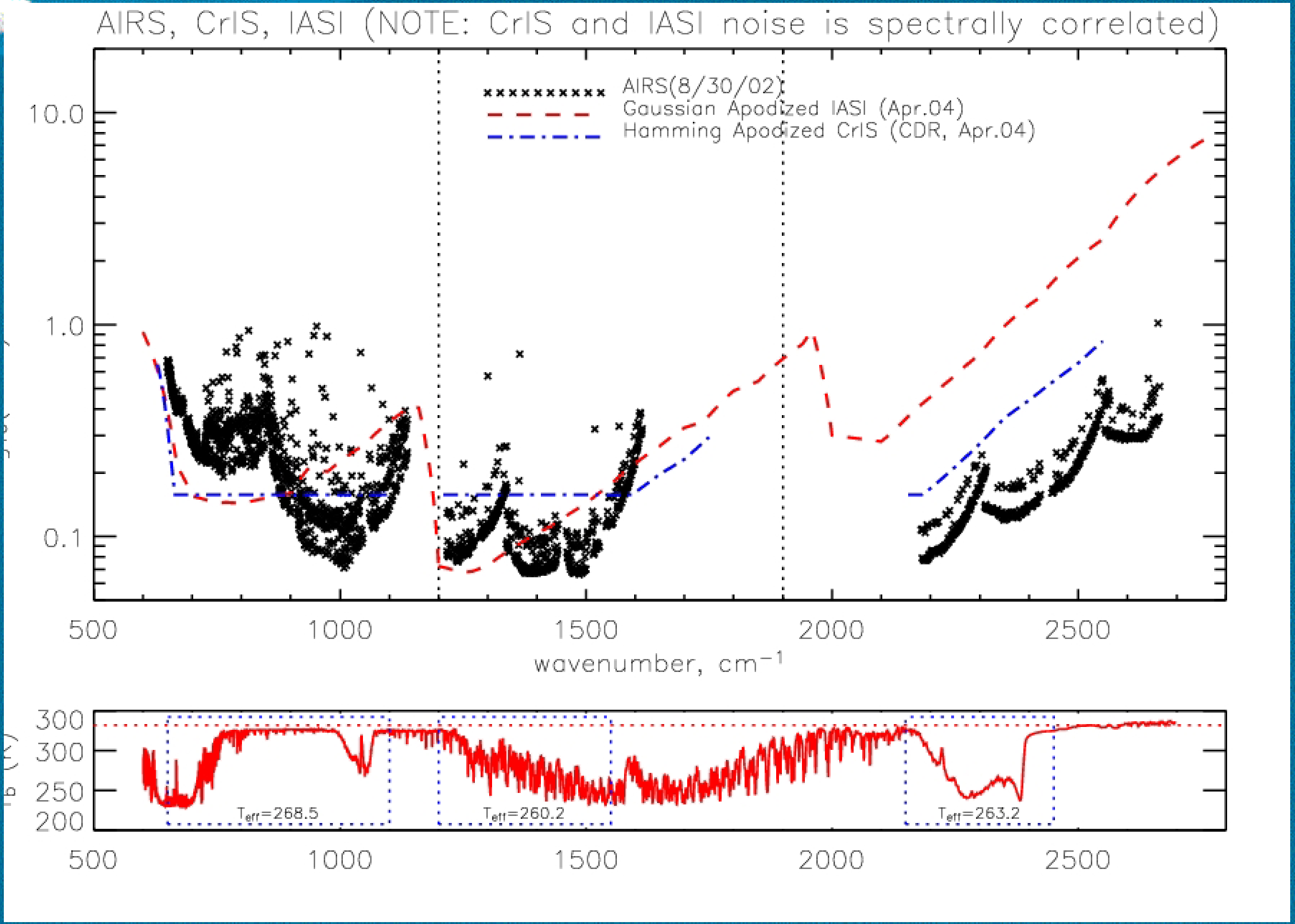
Use cloud-contaminated radiances.

AIRS, CRIS and IASI comparisons

AIRS, CrIS, IASI



NOISE Spectra



CO poor performance by current configuration of CrIS