Application of PCA to IASI:
An NWP Perspective

Andrew Collard, ECMWF

Acknowledgements to Tony McNally, Jean-Noel Thépaut

Overview

- Introduction
  - Why use PCA/RR?
- Expected performance of PCA/RR for IASI
- Assimilation of RR: AIRS Experience
  - Assimilation of “Normal” Radiances
  - Assimilation of Reconstructed Radiances
- Conclusions

Why Use PCA/RR?

Why is data compression important?

- Very large data volumes need to be communicated in near-real time (e.g., EUMETSAT to NWP centres)
- Simulation of spectra (needed for assimilation) is costly
- Data storage

Expected Performance of PCA/RR for IASI?

Spectral data compression with PCA*

The complete AIRS spectrum can be compressed using a truncated principal component analysis (e.g. 200PCAs v 2300 rads)

Leading eigenvectors (200,say) of covariance of spectra from (large) training set

Mean spectrum

$$\mathbf{p} = \mathbf{V}^T(\mathbf{y} - \overline{\mathbf{y}})$$

Coefficients

Original Spectrum

- To use PCs in assimilation requires an efficient RT model to calculate PCs directly
- PCs are more difficult to interpret physically than radiances

N.B. This is usually performed in noise-normalised radiance space

This allows data to be transported efficiently

***Principal Component Analysis**

Spectral data compression and de-noising

The complete AIRS spectrum can be compressed using a truncated principal component analysis (e.g. 200PCAs v 2300 rads)

Leading eigenvectors (200,say) of covariance of spectra from (large) training set

Mean spectrum

Reconstructed spectrum

$$\mathbf{p} = \mathbf{V}^{T}(\mathbf{y} - \bar{\mathbf{y}})$$

$$\mathbf{y}_{R} = \bar{\mathbf{y}} + \mathbf{V}\mathbf{p}$$

Coefficients

Original Spectrum

N.B. This is usually performed in noise-normalised radiance space

Each reconstructed channel is a linear combination of all the original channels and the data is significantly de-noised.

If *N* PCs are used all the information is contained in *N* reconstructed channels (theoretically)

Transformation of Instrument Noise

Error in observation, $\boldsymbol{\varepsilon}_y$, is transformed by:

$$\boldsymbol{\varepsilon}_{\mathbf{y}_R} = \mathbf{V}\mathbf{V}^{\mathbf{T}}\boldsymbol{\varepsilon}_{\mathbf{y}} + \boldsymbol{\varepsilon}_{\mathbf{R}}$$

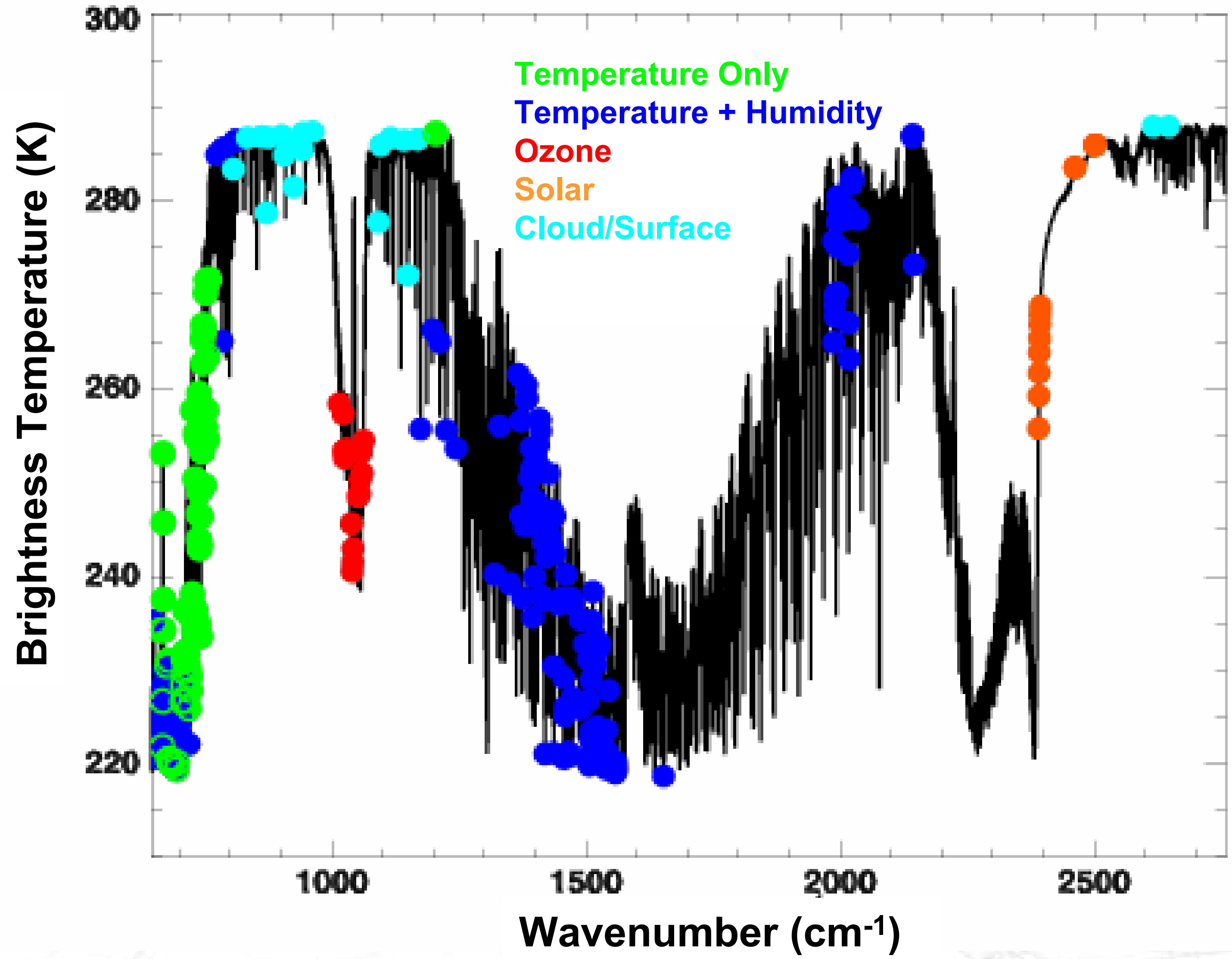
Where $\boldsymbol{\varepsilon}_R$ is the reconstruction error.

∴ Observation error covariance , $\mathbf{O}_R$, of reconstructed radiances:

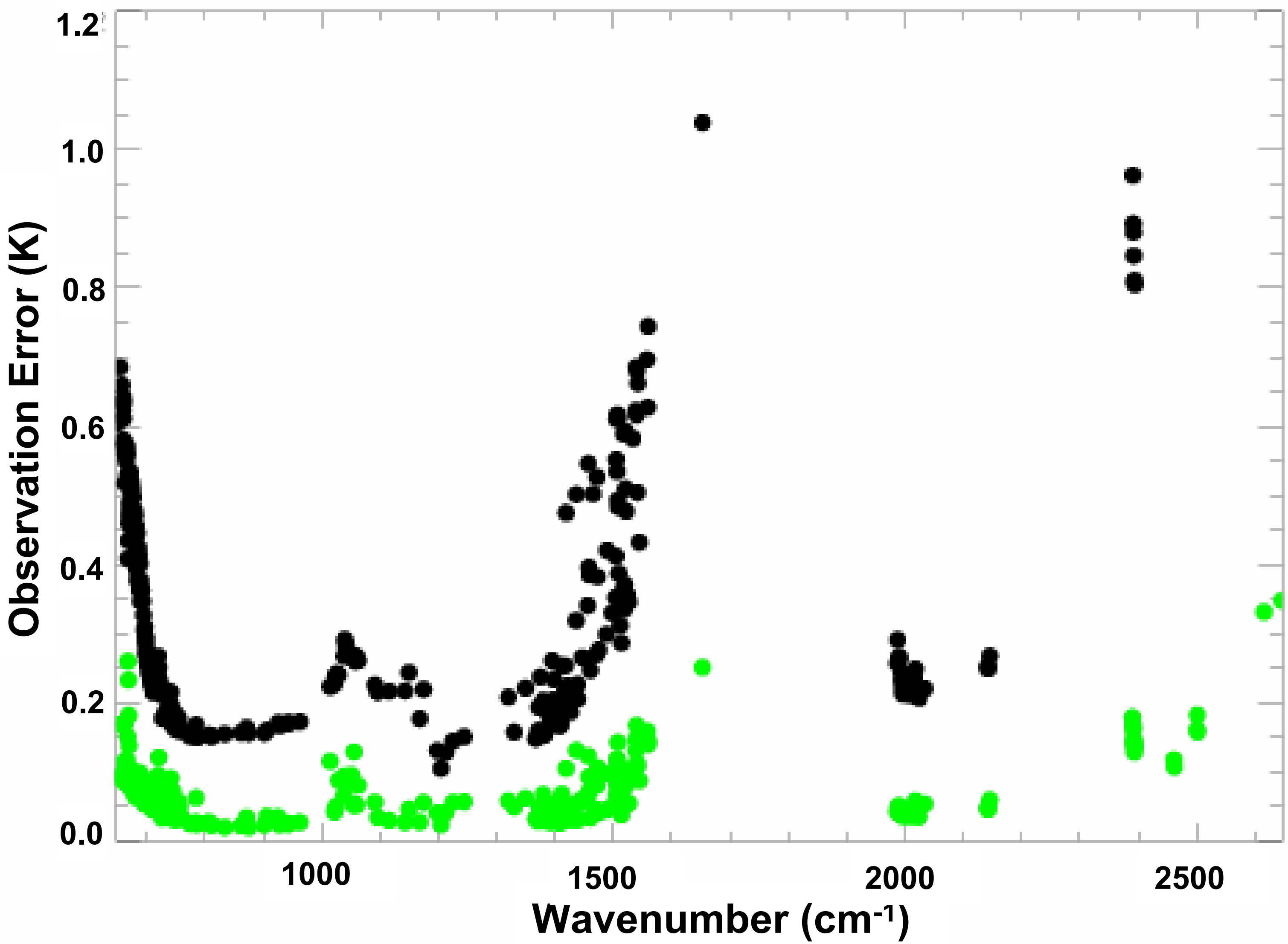
$$\mathbf{O}_{\mathbf{R}} = \mathbf{V}\mathbf{V}^{\mathbf{T}}\mathbf{O}\mathbf{V}^{\mathbf{T}}\mathbf{V} + \mathbf{F}_{\mathbf{R}}$$

In addition, the ***forward model term*** is not modified by reconstruction.

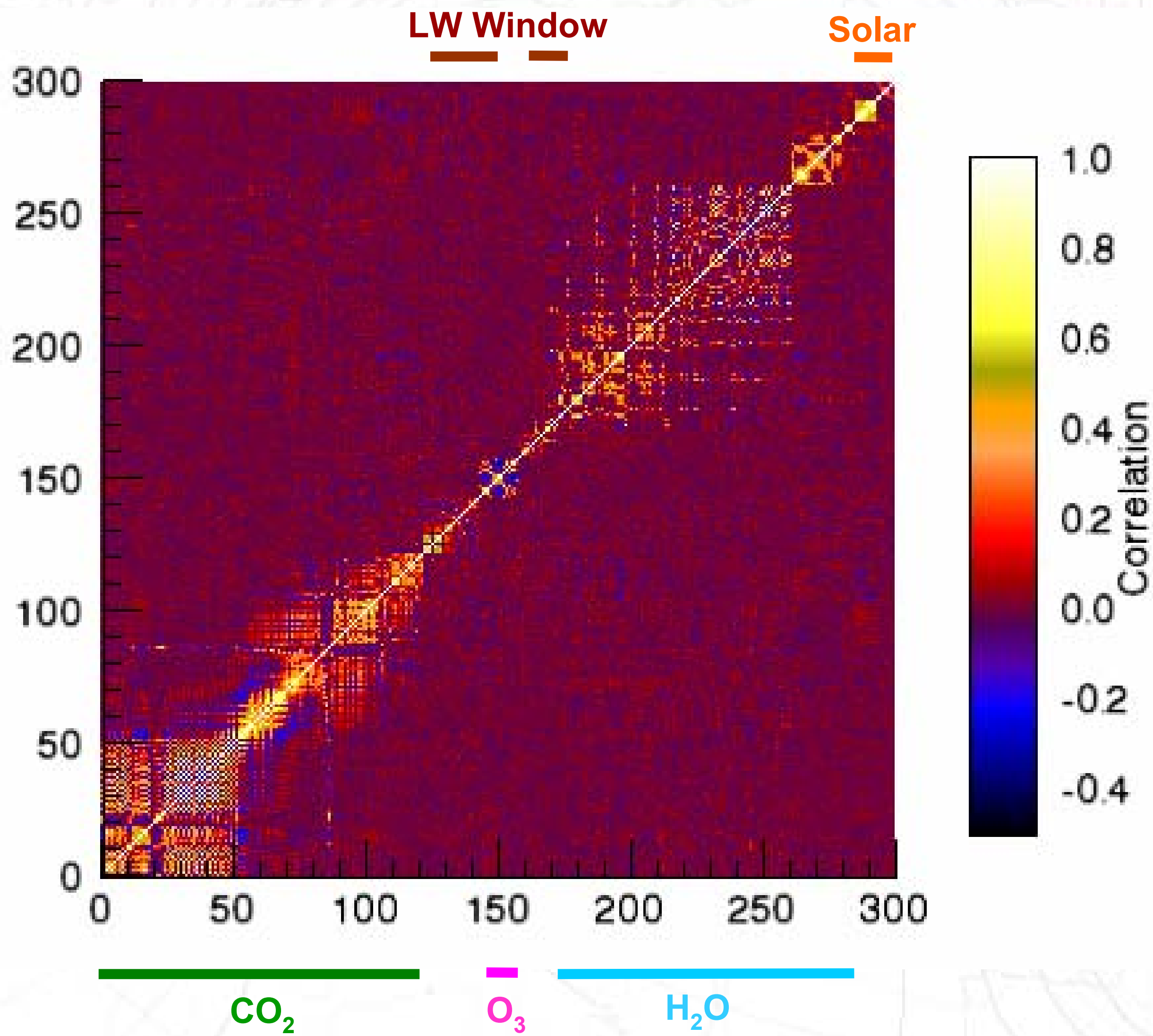
300 Channels for IASI



Noise Reduction – 300 Channels



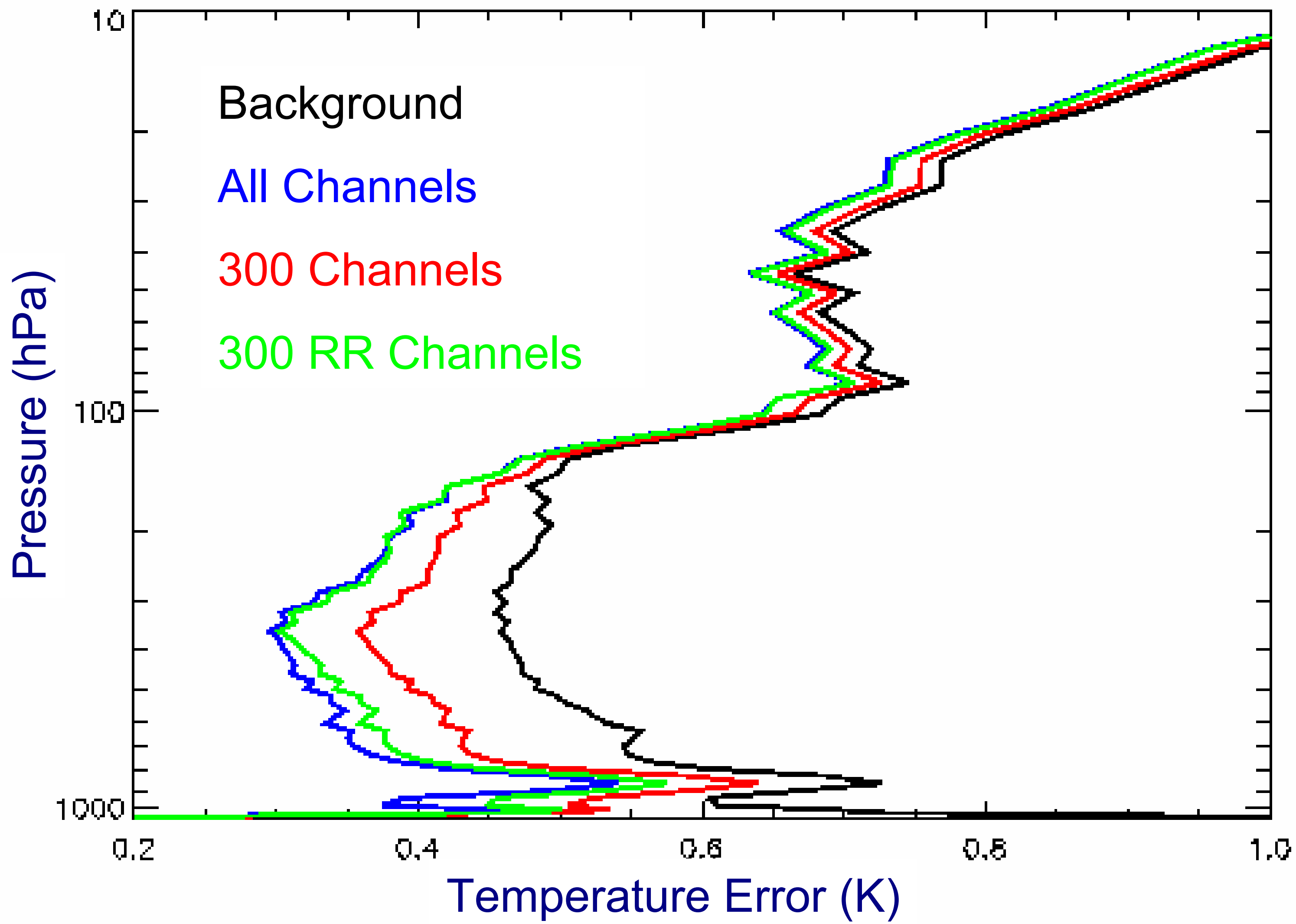
IASI RR Error Correlations



500 PCs

300 IASI Channels

Retrieval Performance



Application of PCA to ~~IASI~~ AIRS:

An NWP Perspective

Andrew Collard, ECMWF

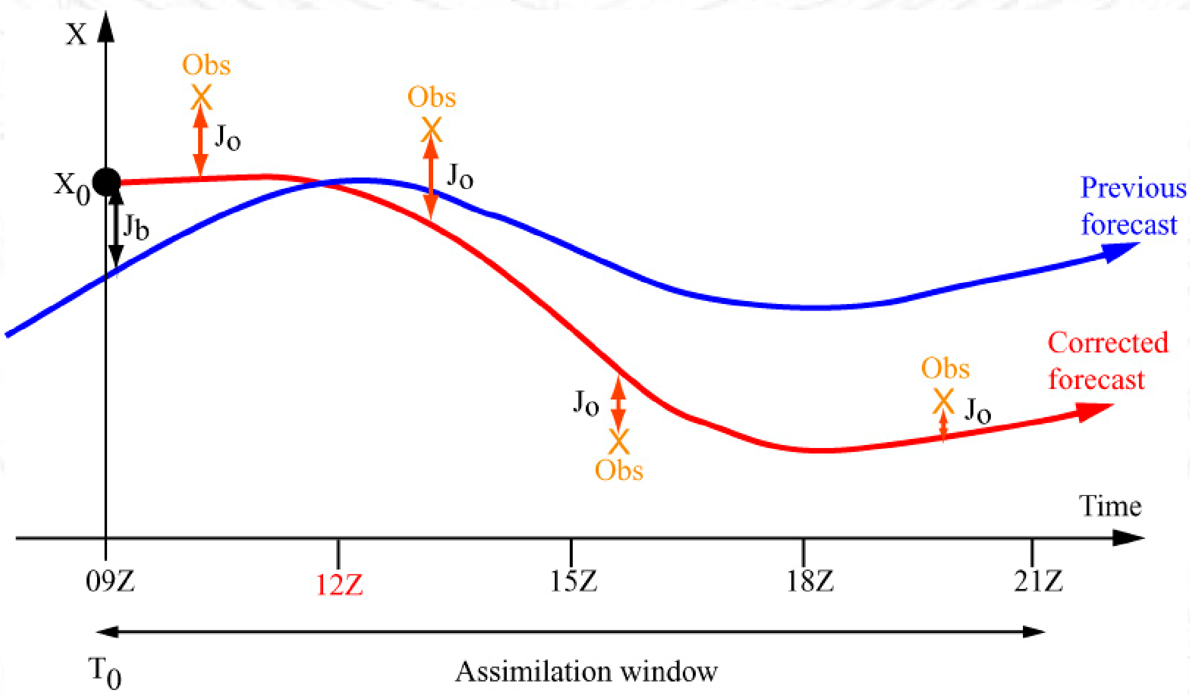
Acknowledgements to Tony McNally, Jean-Noel Thépaut

Experience with AIRS:
Current Operational System

4D-Var Data Assimilation

4-dimensional variational data assimilation is in principle a least-squares fit in 4 dimensions between the predicted state of the atmosphere and the observations.

The adjustment to the predicted state is made at time T_o, which ensures that the analysis state (4-dimensional) is a model trajectory.



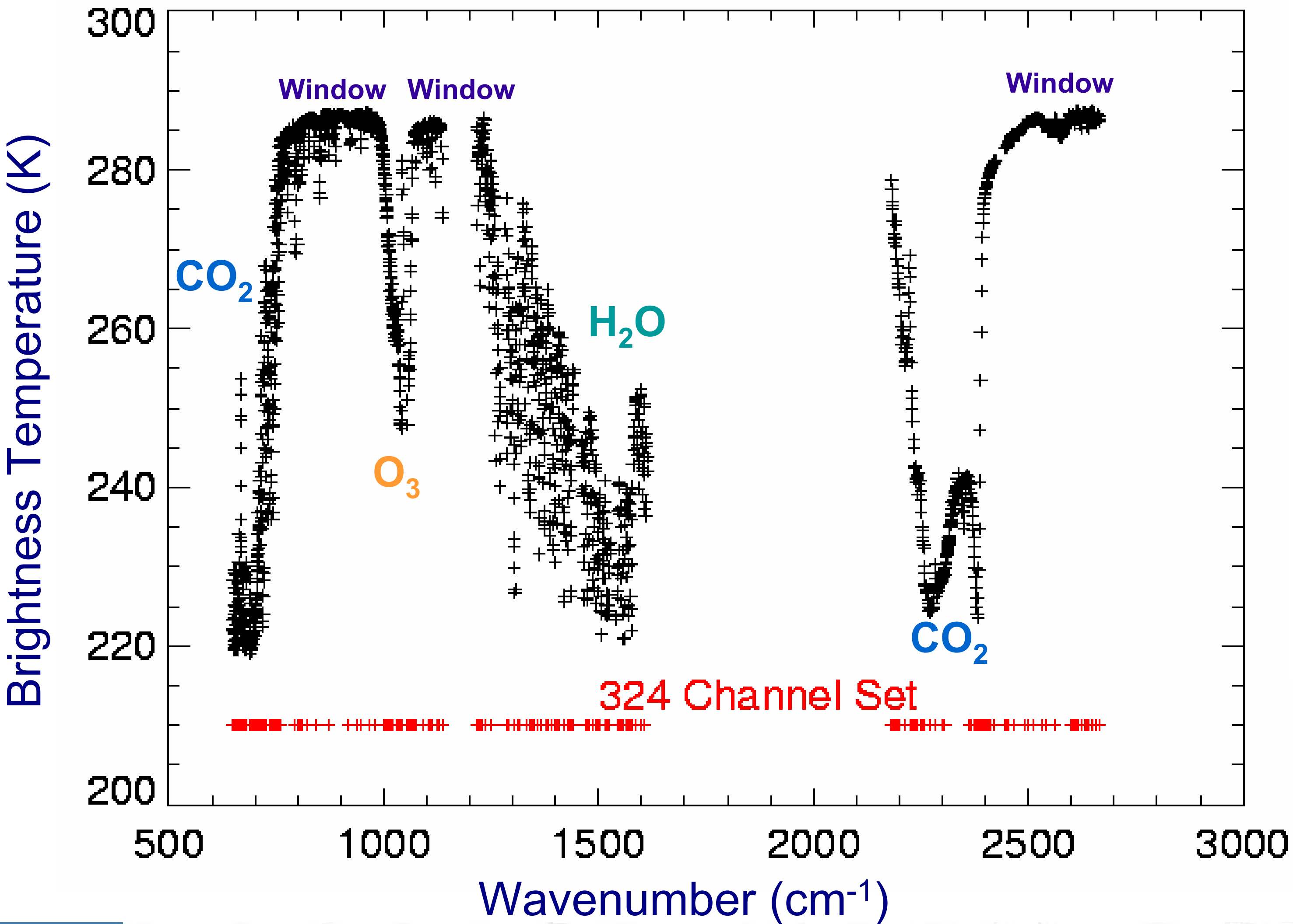
Satellite data assimilated operationally at ECMWF

- 4xAMSU-A (NOAA-15/16/18 + AQUA)
- 3xAMSU-B (NOAA-16/17/18)
- 3 SSMI (F-13/14/15) in clear and rainy conditions
- 1xHIRS (NOAA-17)
- AIRS (AQUA)
- Radiances from 4 GEOS (Met-5, Met-8, GOES-10/12)
- Winds from 4 GEOS (Met-5/8 GOES-10/12) and MODIS/TERRA+AQUA
- Scat winds from QuikSCAT and ERS-2 (Atlantic)
- Wave height from ENVISAT RA2 and ASAR, JASON
- Ozone from SBUV (NOAA 16) and SCIAMACHY (ENVISAT)

29 different satellite sources

Coming soon: SSMIS, radio occultation (GPS),…and IASI!

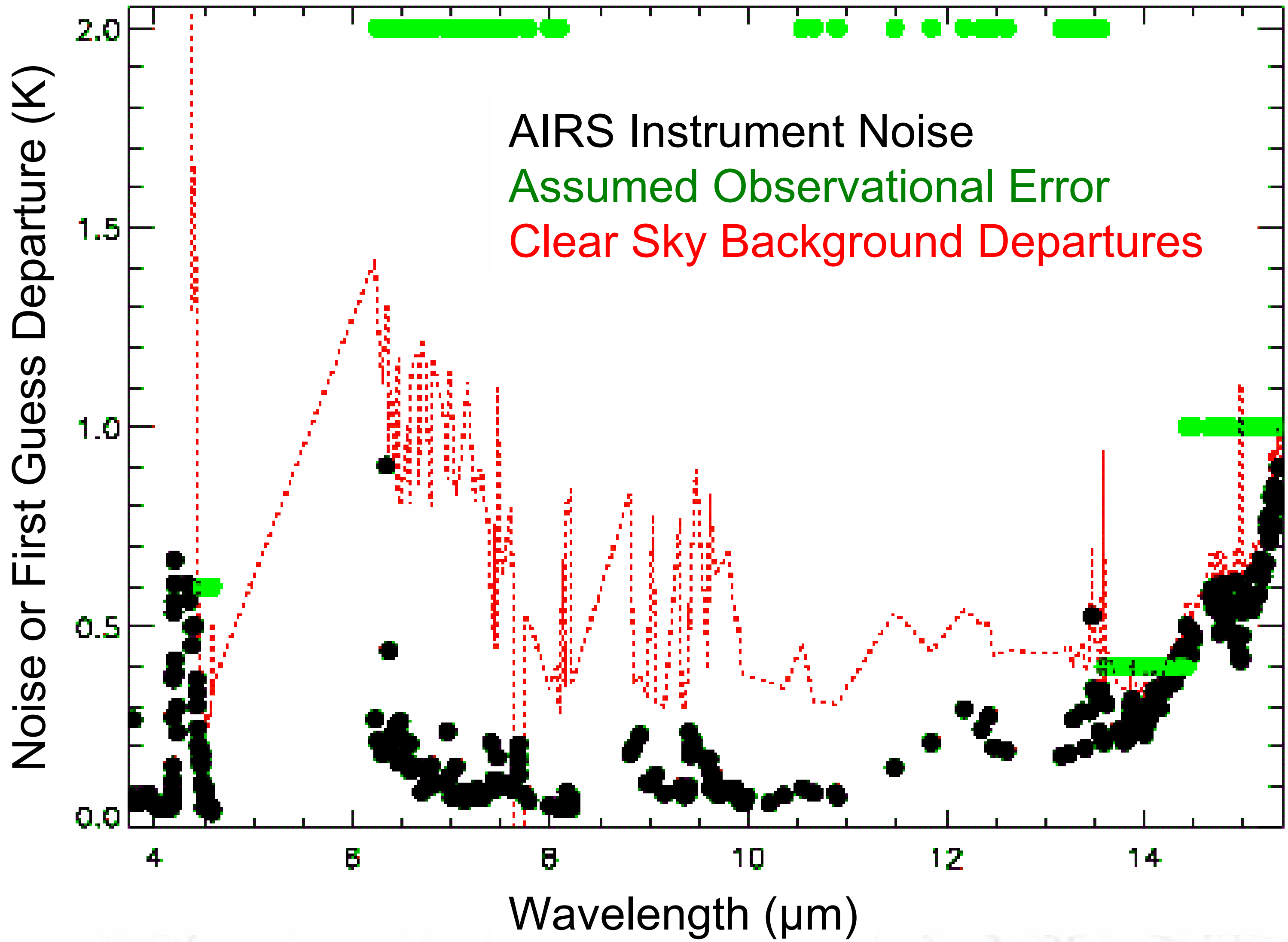
AIRS Spectrum – 324 Channel Subset



- 324 Channels (BUFR dataset)
- One AIRS FOV for every AMSU FOV (there are 9 AIRS FOVS per AMSU FOV)
- Up to 150 channels are assimilated (depending on cloud top height)

N.B. The same channels are supplied in NRT as reconstructed radiances

Assumed Noise for AIRS Assimilation



Optimal vs Non-Optimal Assimilation

- Optimal Assimilation

$$\mathbf{A} = \mathbf{B} - \mathbf{WHB}$$

- Non-Optimal Assimilation (Watts and McNally, 1988)

$$\mathbf{A}' = (\mathbf{I} - \mathbf{WH})\mathbf{B}'(\mathbf{I} - \mathbf{WH})^{\mathrm{T}} + \mathbf{W}\mathbf{O}'\mathbf{W}^{\mathrm{T}}$$

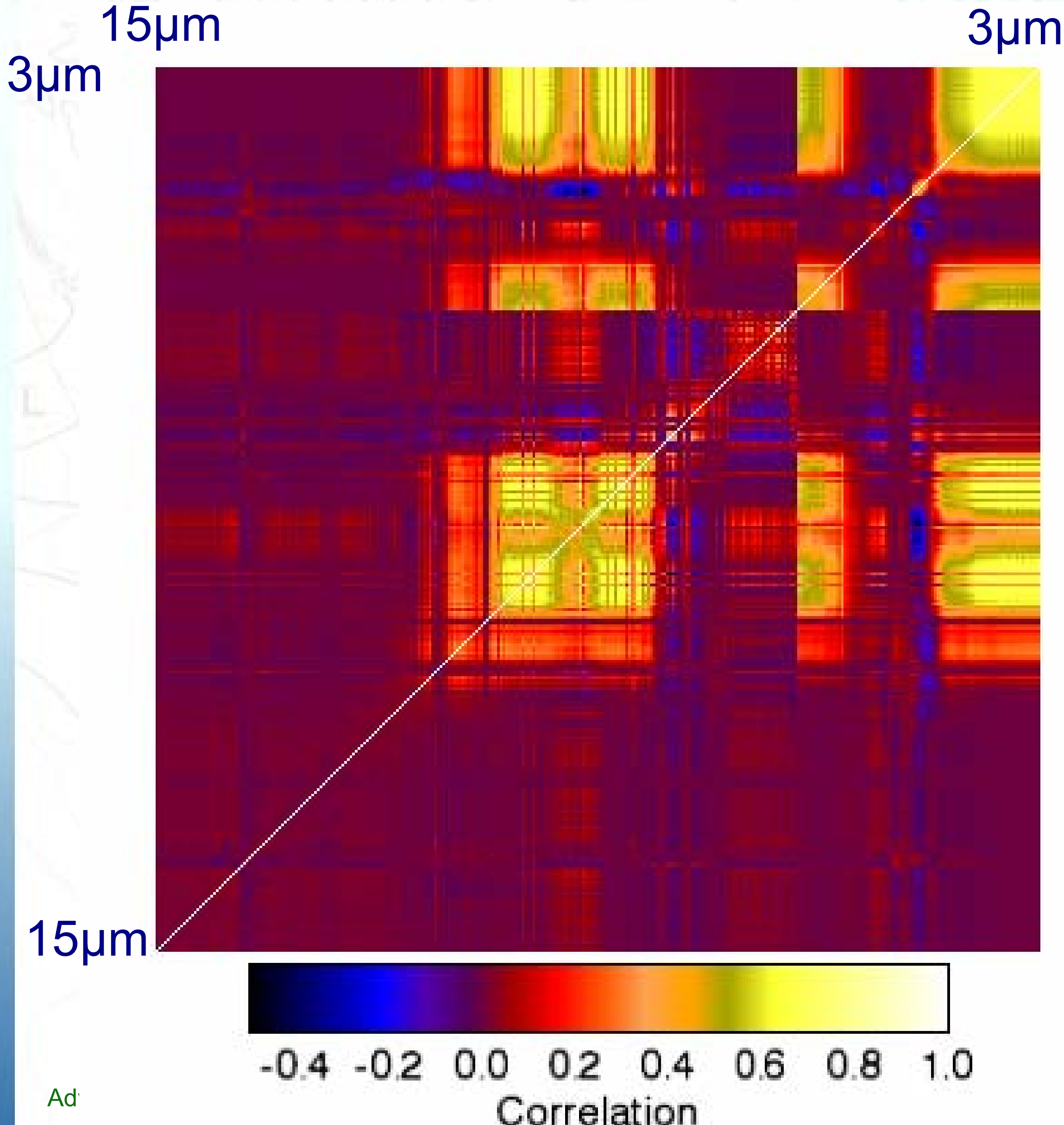
A', **B**' and **O**' are true values for **A**, **B** and **O**

Where

$$\mathbf{W} = \mathbf{BH}^{\mathrm{T}}(\mathbf{HBH}^{\mathrm{T}} + \mathbf{O})^{-1}$$

A=Analysis Error Covariance
B=Background Error Covariance
O=Observation Error Covariance
H=Jacobian

“Observation”* Errors are Correlated



FW Model Error

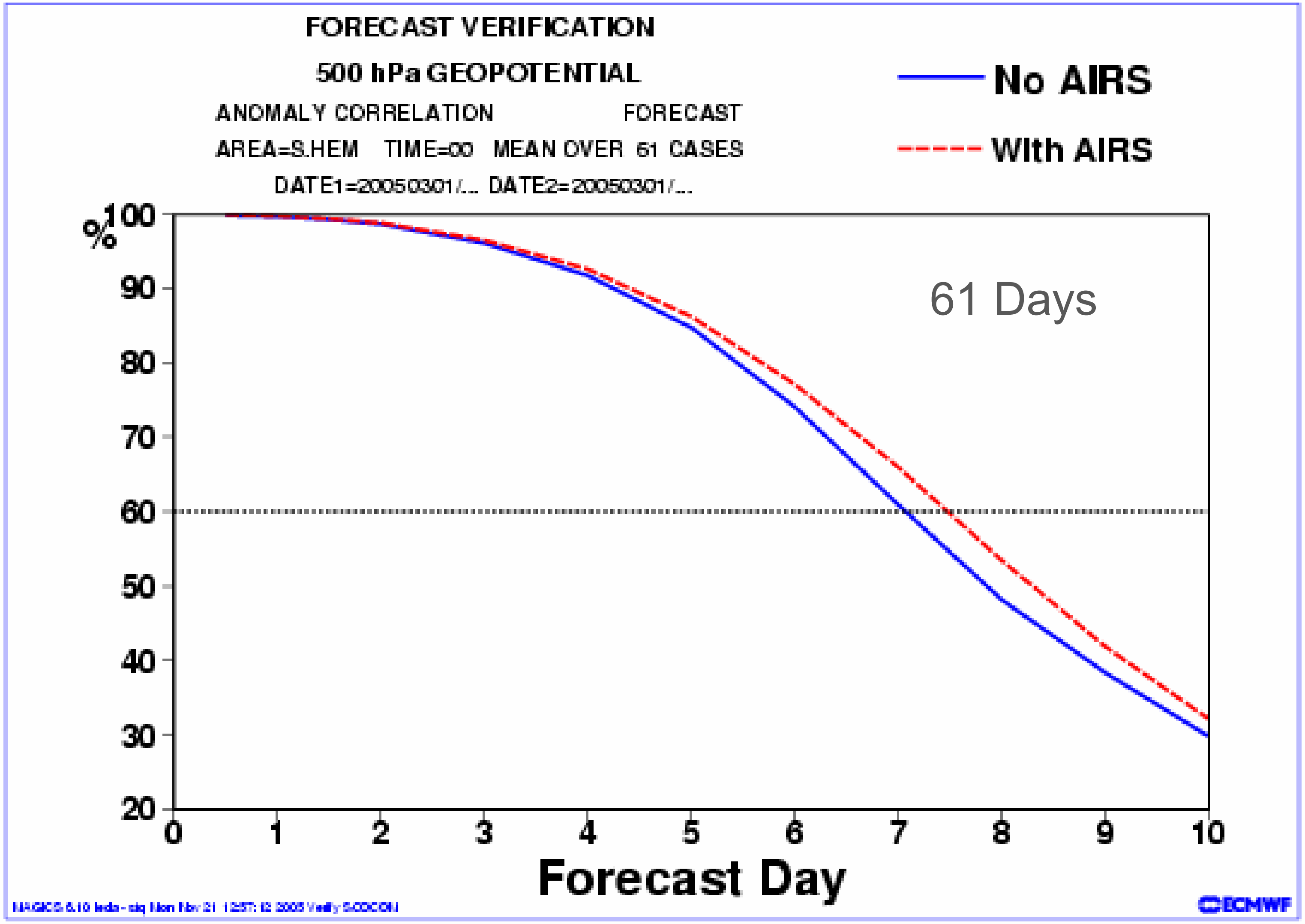
Non-Linearity Error

Cloud/Surface Emissivity Error

Instrument Noise

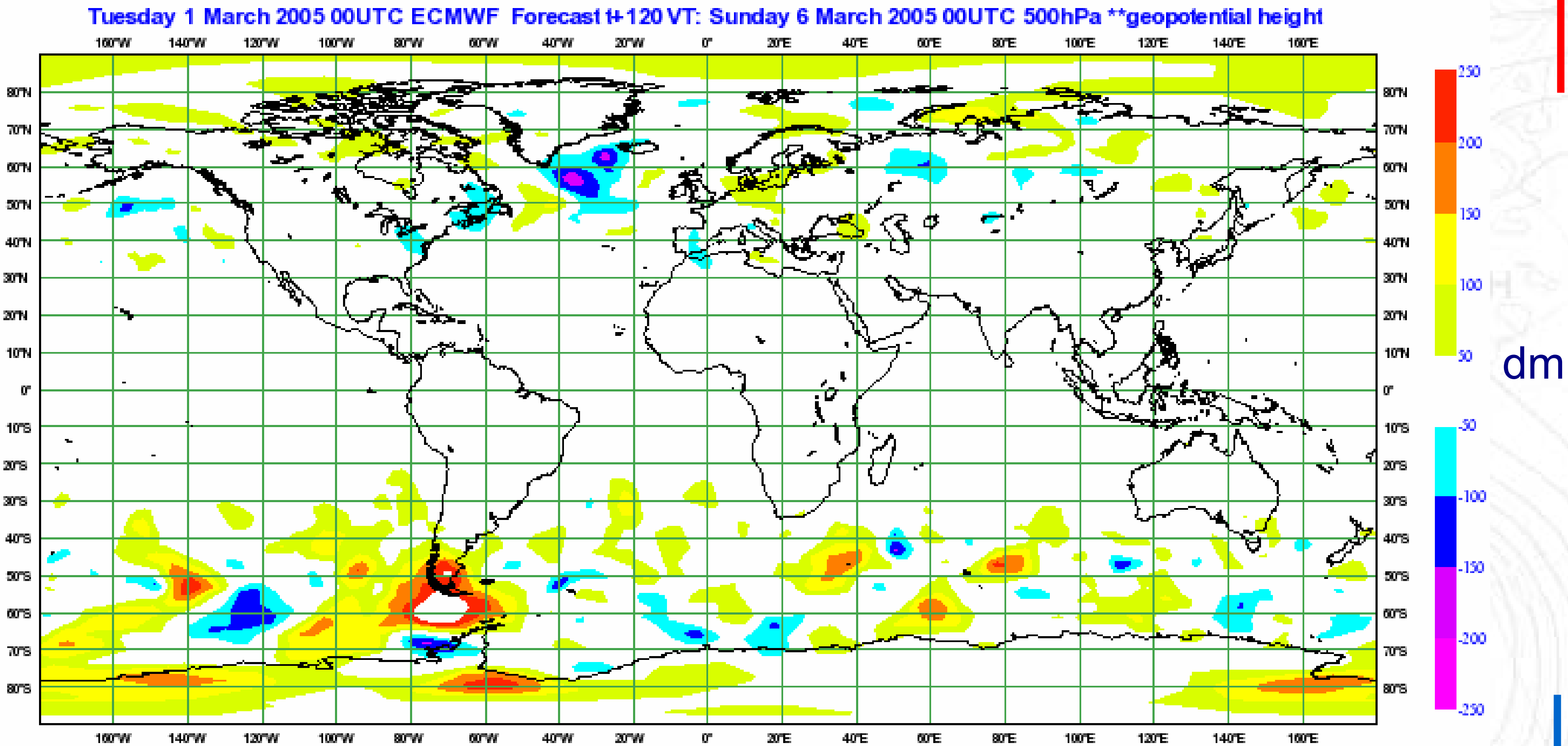
*Instrument plus forward model noise

AIRS Impact up to 10 hours at 7 days



5 Day Forecast Improvements on Adding AIRS

AIRS Improves FC



AIRS Degrades FC

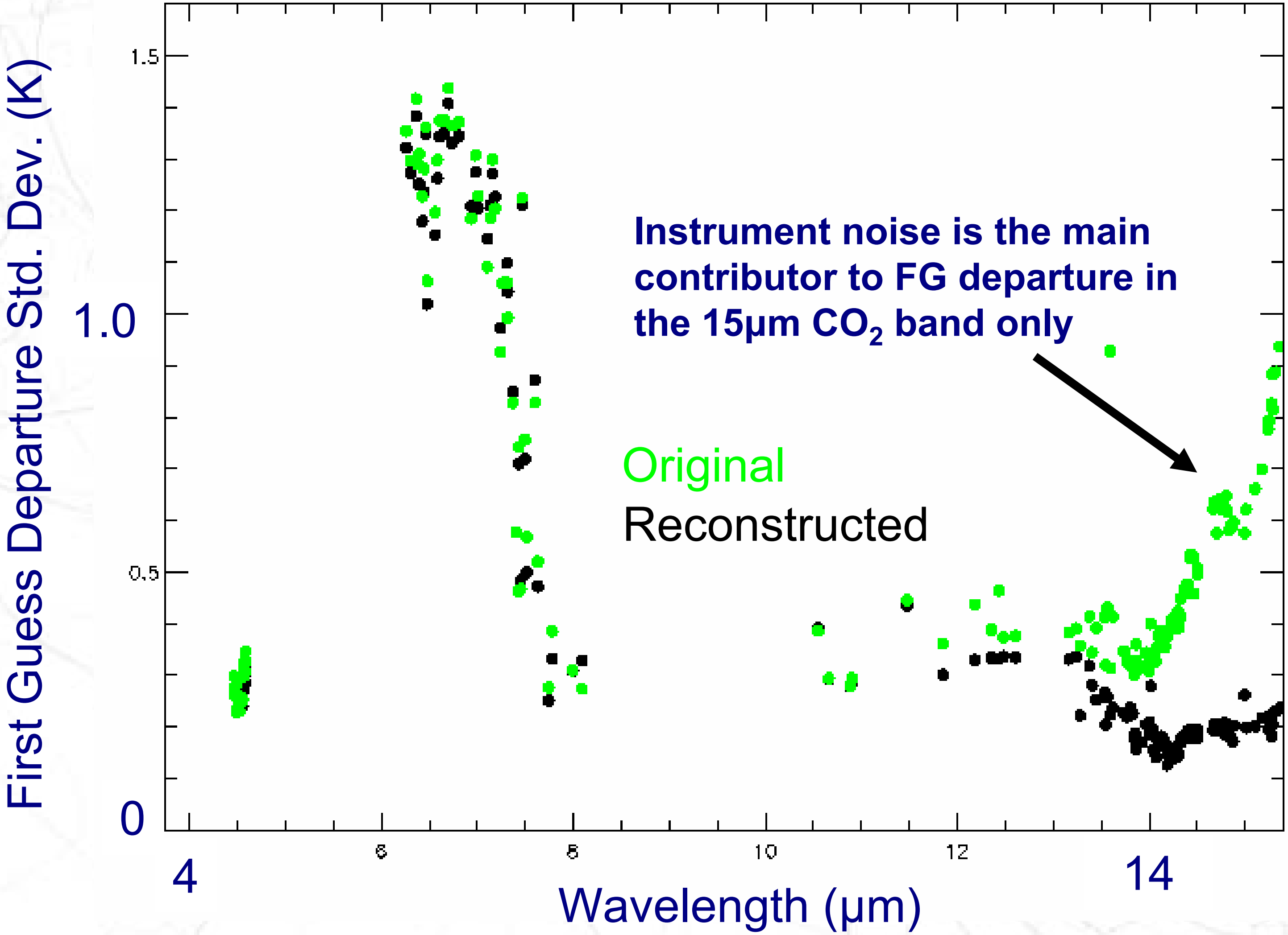
1st March – 30th April 2005

Experience with AIRS:
Reconstructed Radiances

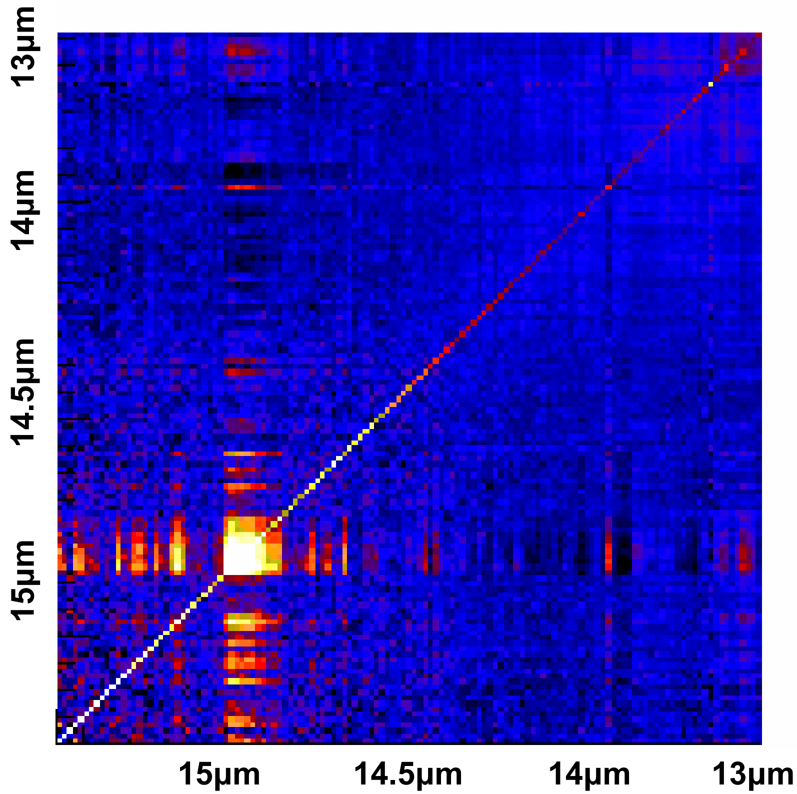
AIRS Reconstructed Radiances

- Data are supplied in near-real time by NOAA/NESDIS in the same format as the “real” radiances.
- The same channels are supplied, except some “popping” channels are missing
- Based on 200 PCs
- QC Flag supplied

First Guess Departures for AIRS are Reduced



A look at Reconstructed Radiances' Errors

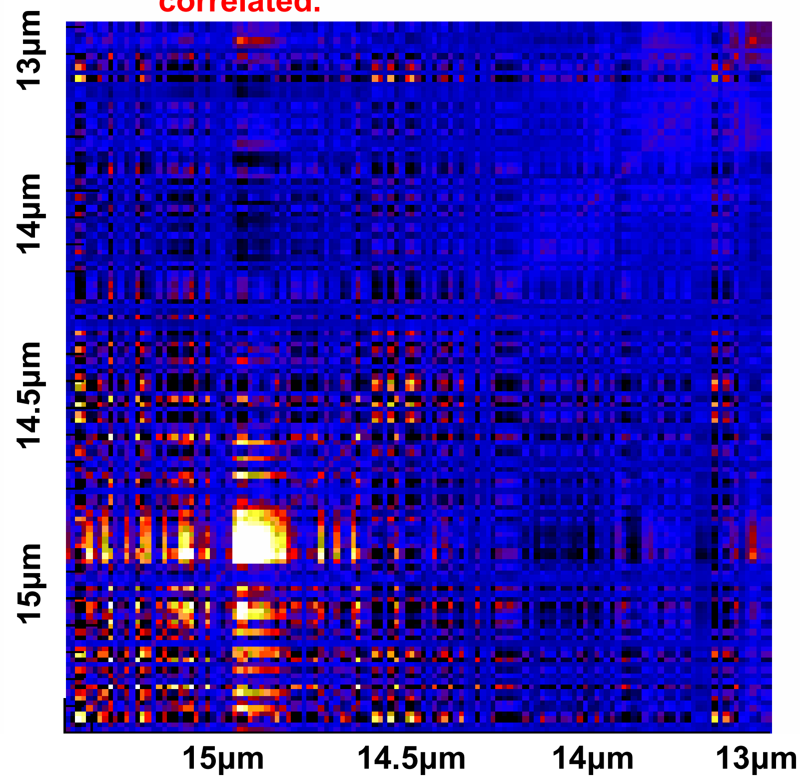


Original Radiances

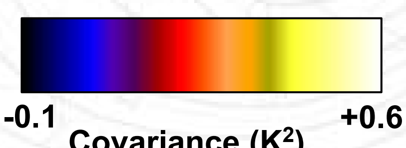
Instrument noise is dominant and diagonal. Correlated noise is from background error

Reconstructed Radiances

Instrument noise is reduced (std. dev. Is approximately halved) but has become correlated.

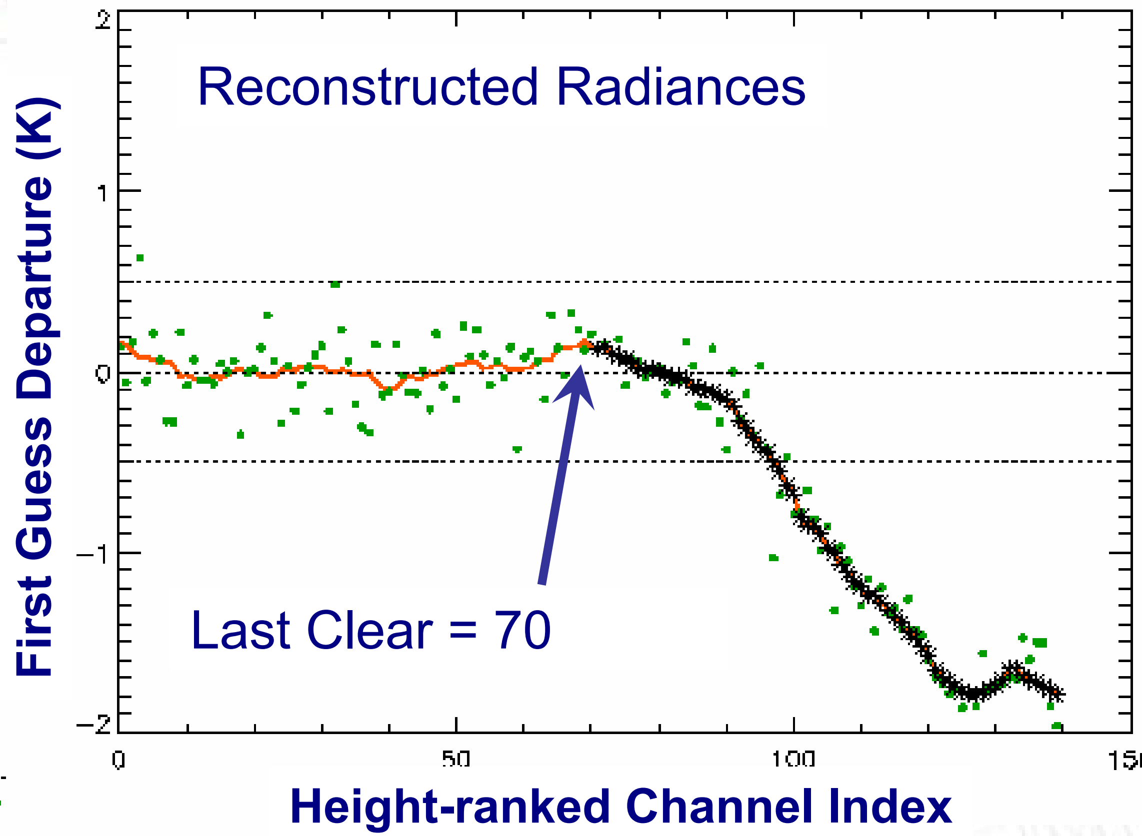


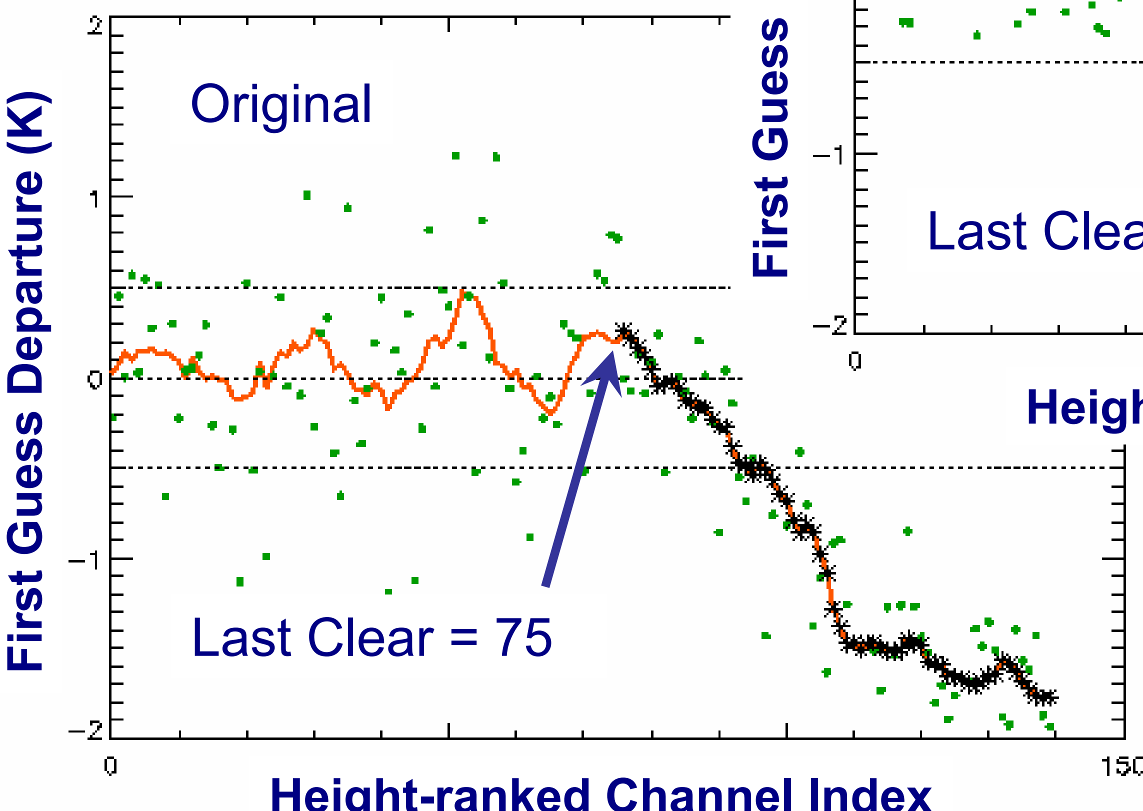
Covariances of background departures for clear observations in 15µm CO_2 band



Improvements in Cloud Detection

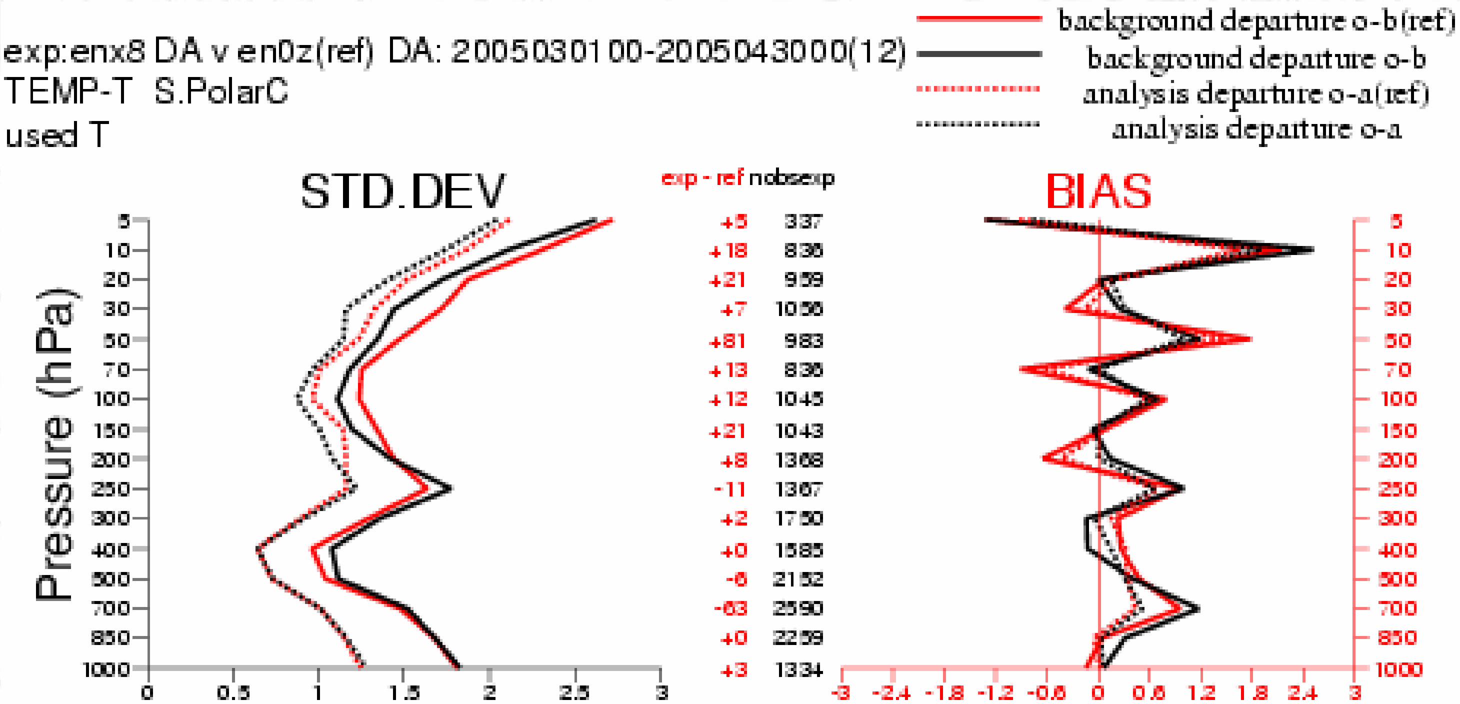
ECMWF Scheme:
- Ranks Channels Height
- Applies low-pass filter
- Tests for non-zero gradient





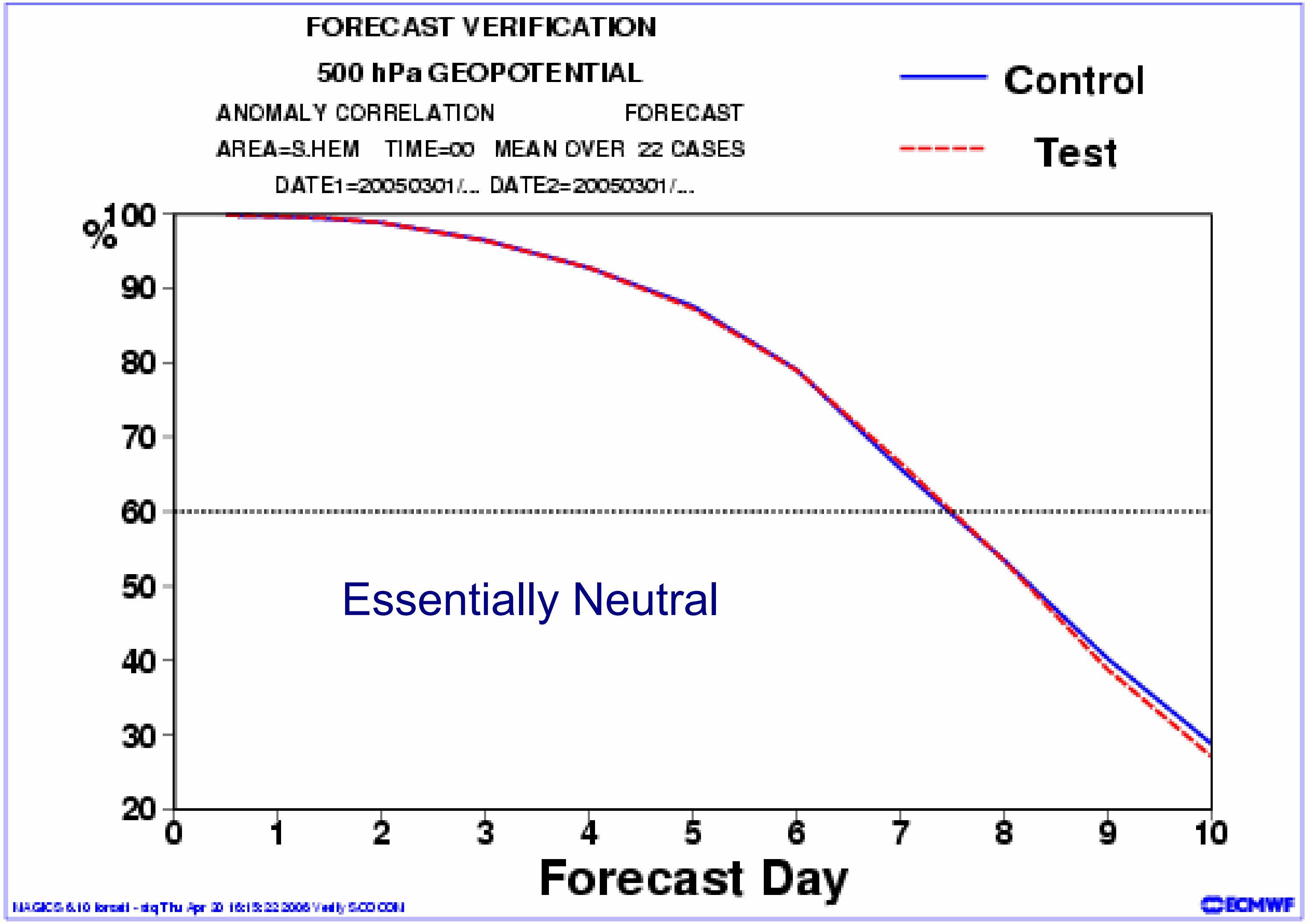
< 2% of Channels are flagged differently for RR vs Normal radiances

Improvements to Antarctic Stratosphere

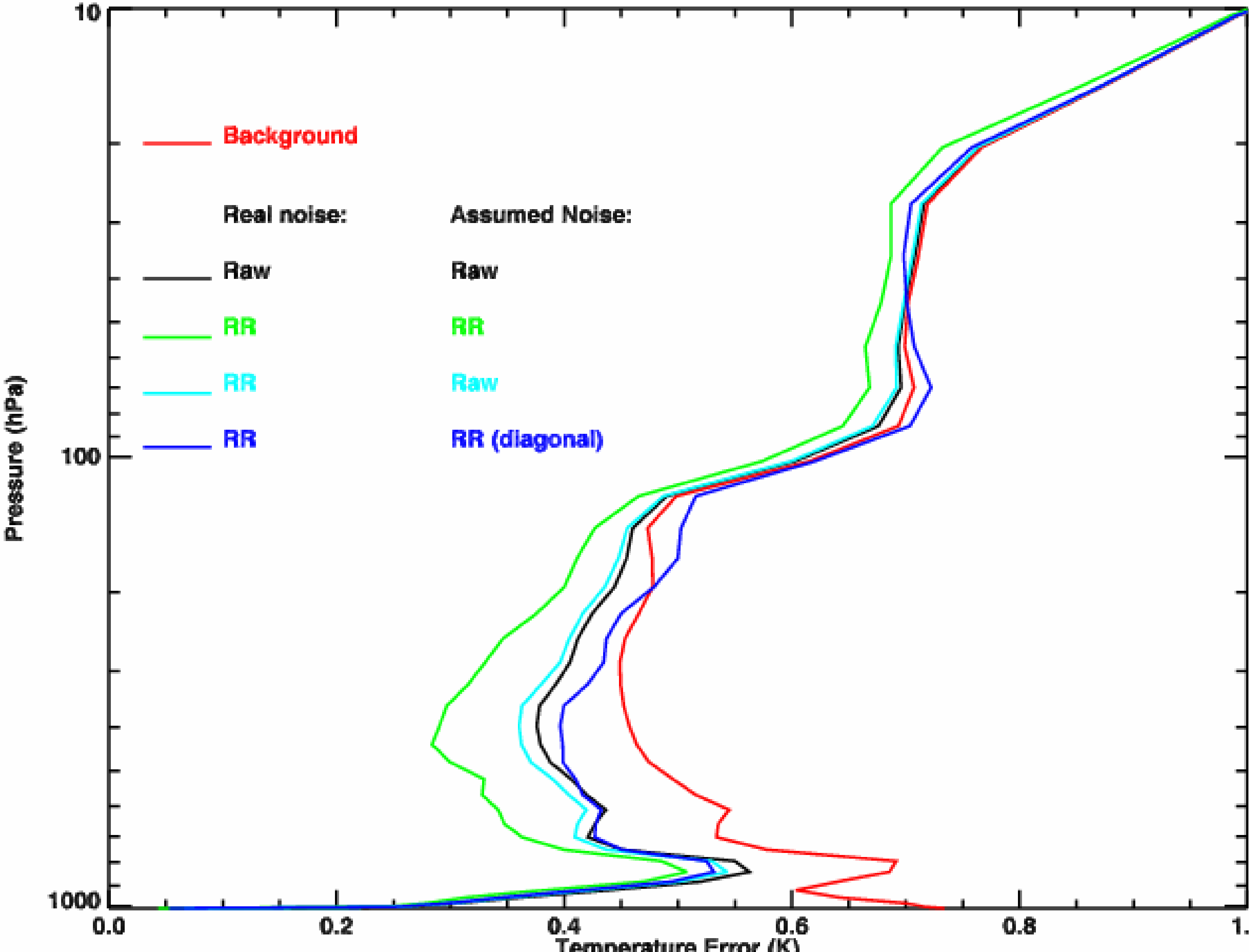


"Stratospheric Oscillation" in comparison
to Antarctic radiosondes is greatly reduced
on moving to reconstructed radiances

Forecast Impact of Reconstructed Radiances



Assimilating Reconstructed Radiances – Linear Theory



So why not use the RR data more aggressively?

- More aggressive errors have been tried:
  - Reduction of existing diagonal errors
  - “Bottom up” construction of full covariances
  - Hollingsworth-Lönnberg approach

- Results have been neutral or negative with respect to:
  - Fit to other observations
  - Performance of short/medium range forecasts

Next steps

- Other error sources need to be investigated:
  - Bias
  - Spatially correlated error
  - Representivity error
  - ????

Conclusions

- The reconstructed radiances method has the potential to be particularly useful for the efficient representation of the information in IASI and AIRS.
- The noise smoothing properties of the reconstructed radiances method shows some positive benefits on assimilation into a high-resolution NWP system.
- However, the assimilation of reconstructed radiances has yet to yield positive impact on ***forecast skill*** (relative to "normal" radiances).
- Work on the construction of a suitable (non-optimal) error covariance continues.