Satellite data assimilation for NWP: II

Sean Healy
European Centre for Medium-range Weather Forecasts (ECMWF)

with contributions from many ECMWF colleagues

Special thanks to: Jean-Noël Thépaut (based on his lecture)
Tony McNally, Niels Bormann, Stephen English, Peter Bauer, Alan Geer …

Outline

1. **Review of concepts from previous lecture.**
   - We spend a lot of our time thinking about errors/limitations of satellite data.
2. **Background errors and vertical resolution**
   - The concept of the observation null space
3. **Systematic biases and bias correction**
4. **Ambiguity in radiance observations**
5. **Quality control.**
6. **Some “new” observation types**
7. **Summary**

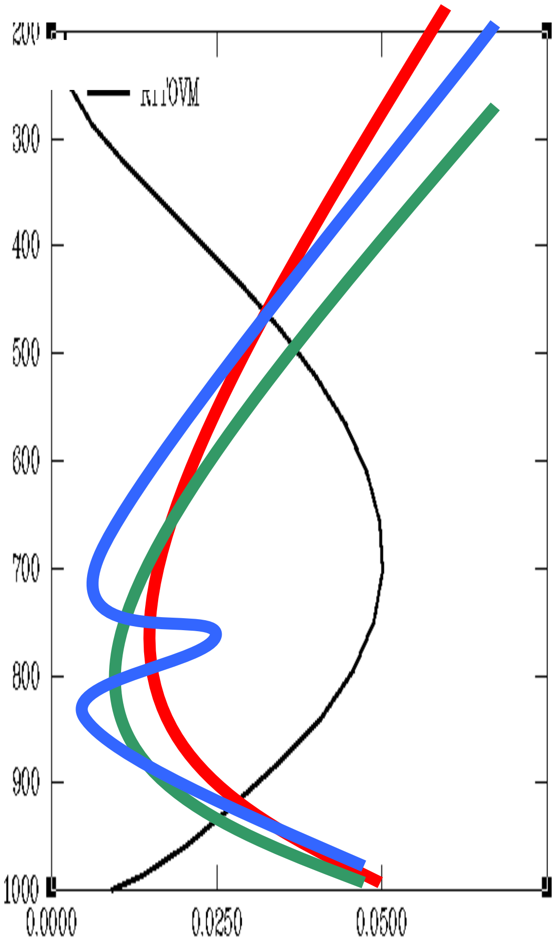
Review of some key concepts

- **Be aware of the difference between the satellite observation and the satellite product/retrieval. (NWP experience in 1980s).**
- **Satellite data are extremely important in NWP.**
- **Data assimilation combines observations and a priori information in an optimal way and is analogous to the retrieval inverse problem.**
- **Passive nadir sounders have the largest impact on NWP forecast skill:**
  - Nadir sounders measure radiance (not T,Q or wind).
  - Sounding radiances are broad vertical averages of the temperature profile (defined by the weighting functions).
  - The retrieval of atmospheric temperature from the radiances is ill-posed and all retrieval algorithms use some sort of prior information.
  - Most NWP centres assimilate raw radiances directly due to their simpler error characteristics. 4DVAR is now widely used (but hybrid techniques have emerged, Ensemble Kalman filters).

2.) Background errors, vertical resolution and null-space.

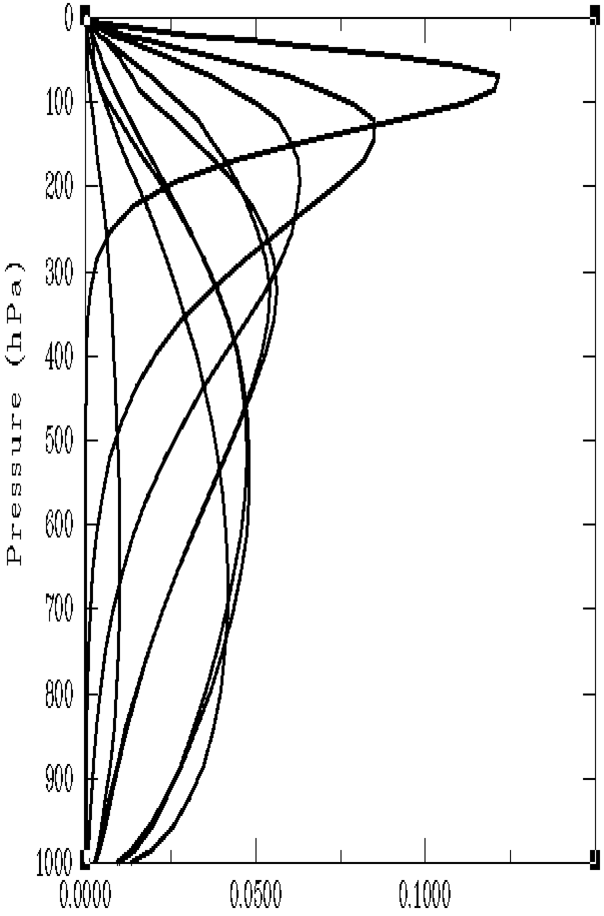
Lecture 1: Satellite radiances have limited vertical resolution

Single channel



Selecting radiation in a number of frequencies / channels improves vertical sampling and resolution

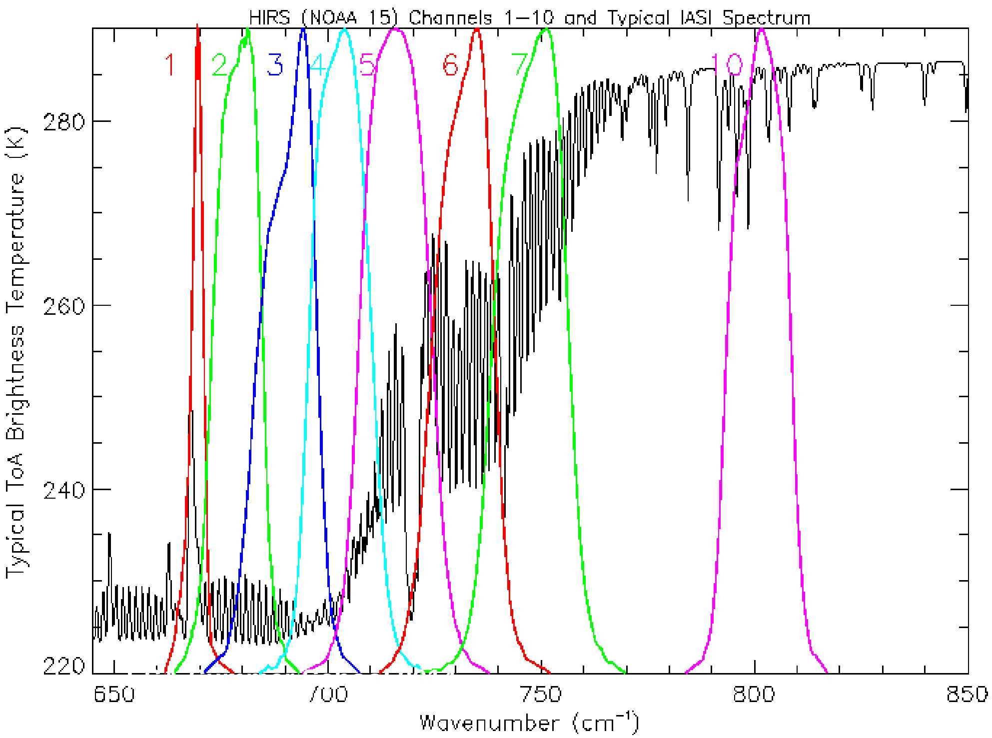
Several channels (e.g. AMSUA)

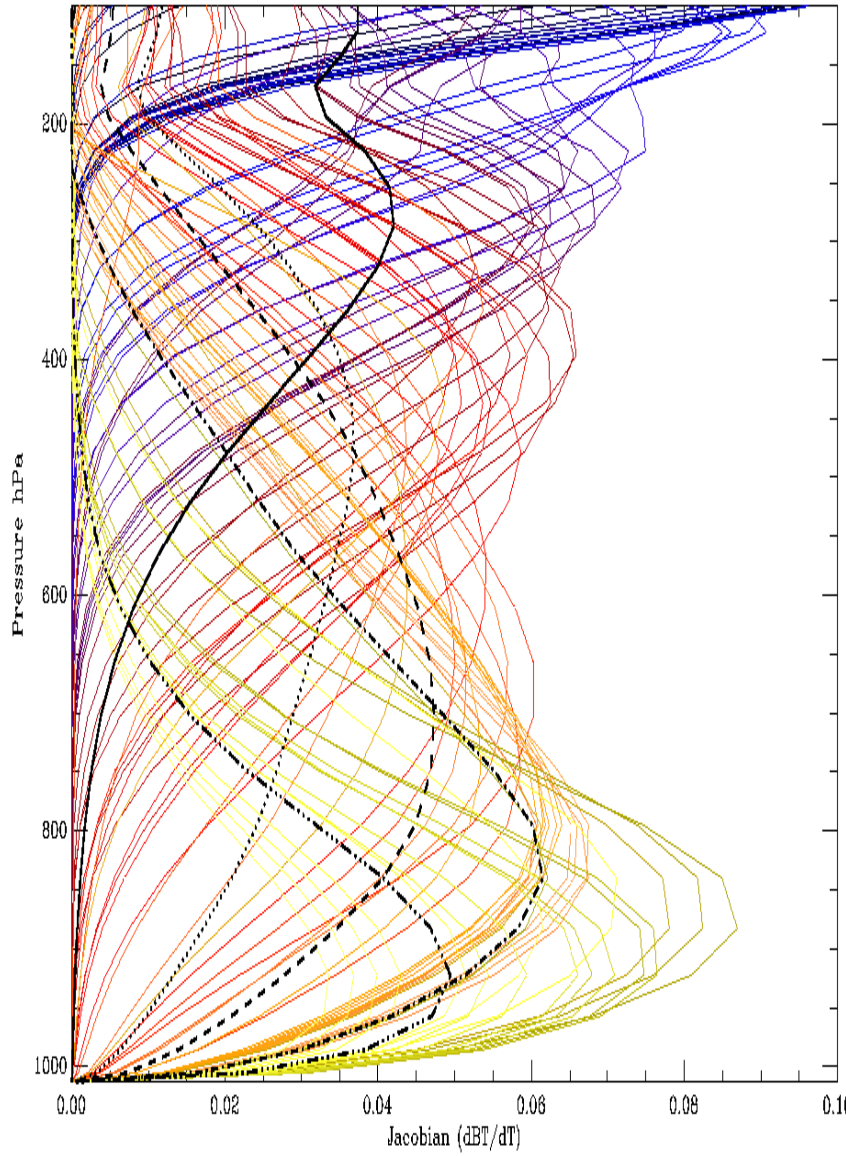


ECMWF

Improving vertical resolution with hyper-spectral IR instruments (AIRS/IASI)

These instruments sample the spectrum extremely finely and thus generate many thousands of channels peaking at different altitudes.
However, vertical resolution still limited by the physics.





Satellite radiances "seeing" and "correcting" background errors

When we minimize a cost function of the form (in 1D / 3D / 4D-VAR)

$$J(\mathbf{x}) = (\mathbf{x} - \mathbf{x_b})^{\mathrm{T}} \mathbf{B}^{-1} (\mathbf{x} - \mathbf{x_b}) + (\mathbf{y_m} - \mathbf{H}[\mathbf{x}])^{\mathrm{T}} \mathbf{R}^{-1} (\mathbf{y_m} - \mathbf{H}[\mathbf{x}])$$

We can think of the adjustment process as radiances observations **correcting errors in the forecast background** to produce an analysis that is closer to the true atmospheric state. For example in the simple linear case...

$$\mathbf{x_a} = \mathbf{x_b} + \underline{\mathbf{B}\mathbf{H}^{\mathrm{T}}(\mathbf{H}\mathbf{B}\mathbf{H}^{\mathrm{T}} + \mathbf{R})^{-1}(\mathbf{y_m} - \mathbf{H}\mathbf{x_b})} \longleftarrow \text{correction term}$$

Because of broad weighting functions the radiances have very little vertical resolution and the **vertical distribution of forecast errors** is crucial to how well they will be "seen" and "corrected" by satellite data in the analysis.

This vertical distribution is communicated to the retrieval / analysis via the **vertical correlations** implicit in the background error covariance matrix **B** (the rows of which are sometimes known as **structure functions**).

The "null space" of a measurement

- We have already noted that infinitely many atmospheric states can produce the same measurement (ill-posed).

$$\mathbf{H}\mathbf{x}_1 = \mathbf{H}\mathbf{x}_2 \sim \mathbf{y}$$
$$\mathbf{H}(\mathbf{x}_1 - \mathbf{x}_2) \sim \mathbf{0}$$

Ok perhaps not = 0, but small compared to the assumed ob. errors.

- So the measurement can't distinguish between these two states. (Null-space: look at a SVD of H matrix.)

- What happens if the NWP forecast error, $\epsilon \propto (\mathbf{x}_1 - \mathbf{x}_2)$?

- Bad luck. The measurement can't correct this error.

Simple example

- **Again I measure the scalar**

$$y = x_1 + x_2$$

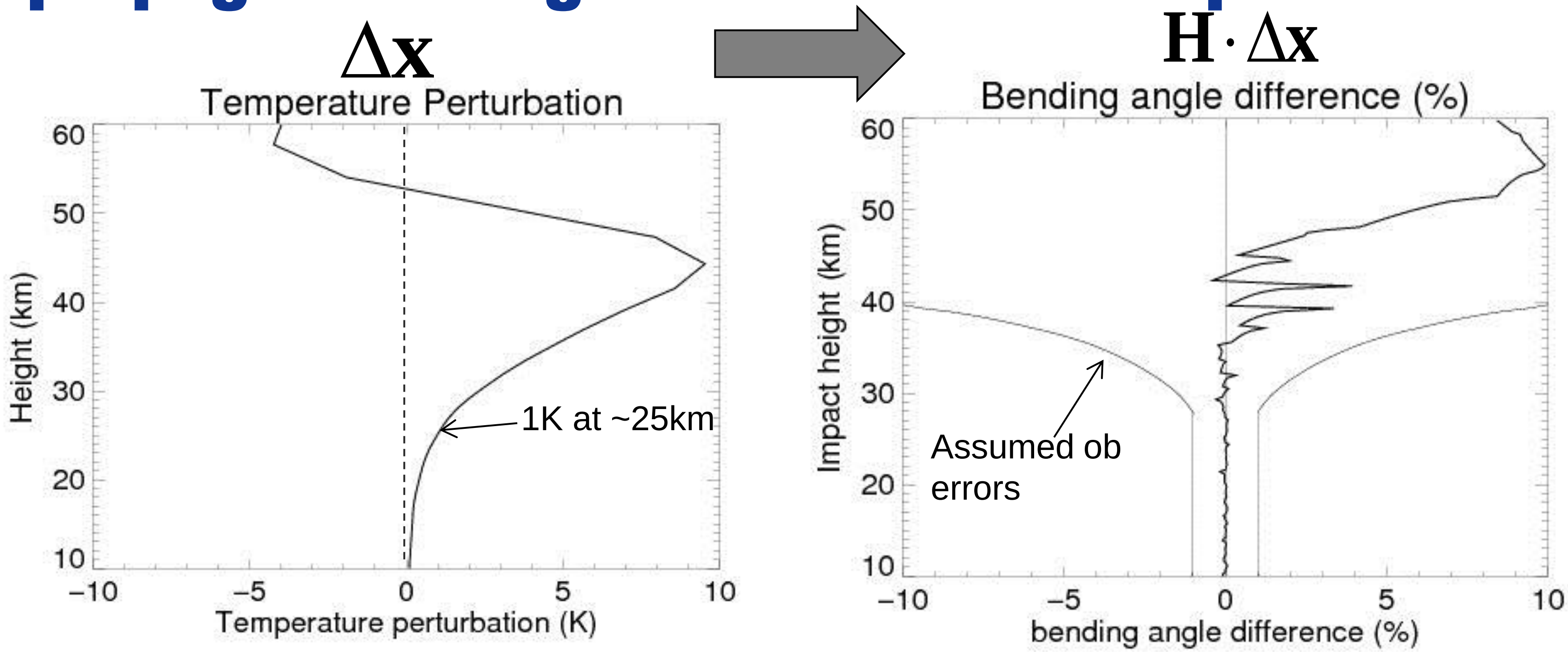
- **I have model that provides me with an estimate**

$$x_b = \begin{pmatrix} x_1^b \\ x_2^b \end{pmatrix} = \begin{pmatrix} x_1^t + \varepsilon_1 \\ x_2^t + \varepsilon_2 \end{pmatrix}$$

- **But (unfortunately) the background error is** $\varepsilon_1 \approx -\varepsilon_2$

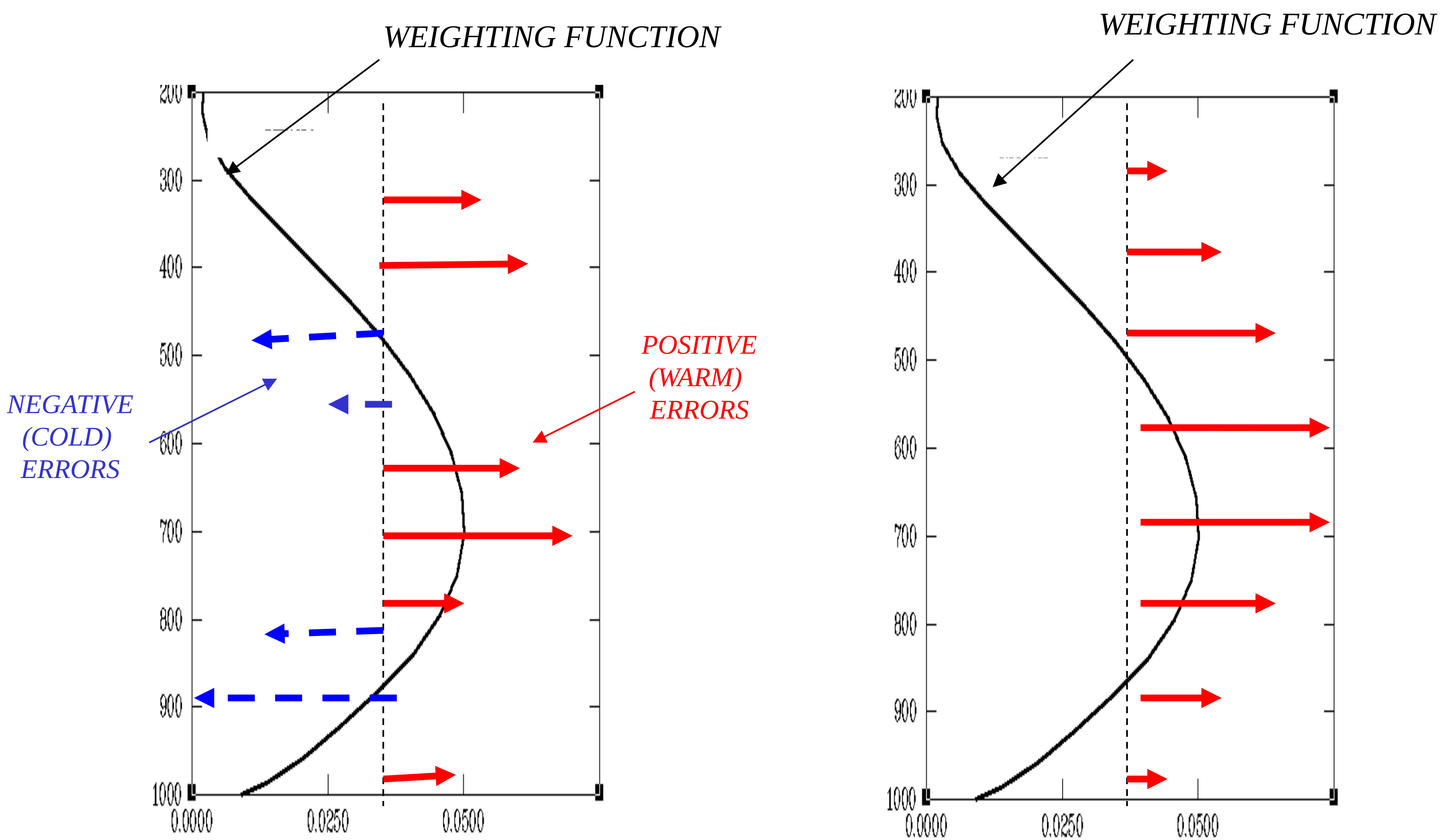
- **The background errors could be huge, but they are in the measurement null space, so the measurement can't constrain them. You have to know the null space!**

Real example: GPSRO null space – how does the temperature difference at the S.Pole propagate through the observation operator



The null space arises because the measurements are sensitive to density as function of height (~P(z)/T(z)). *A priori* information is required to split this into T(z) and P(z). We can define a temperature perturbation ΔT(z)~k*exp(z/H) which is in the GPSRO null space. *Therefore, if the model background contains a bias of this form, the measurement can't see or correct it*.

Correcting errors in the background



"Difficult" to correct

"Easy" to correct

ECMWF

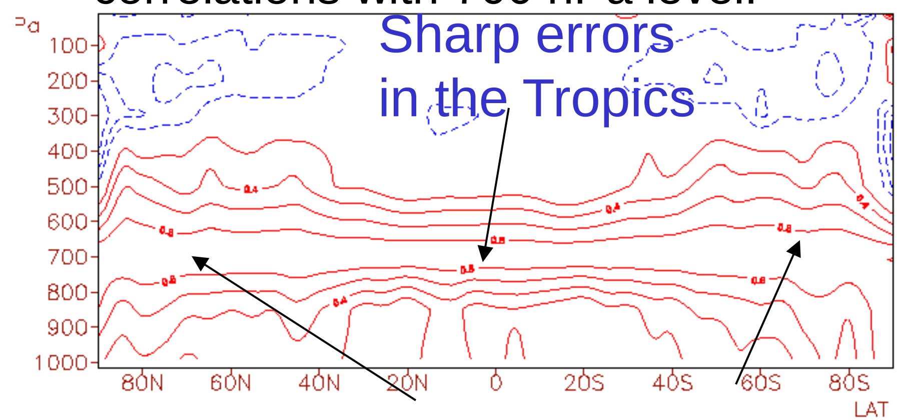
Estimating background error correlations

If the **background errors are mis-specified** in the retrieval / analysis this can lead to a complete mis-interpretation of the radiance information and badly damage the analysis, possibly producing an analysis with **larger errors than the background state** **!**

Thus accurate estimation of **B** is crucial:

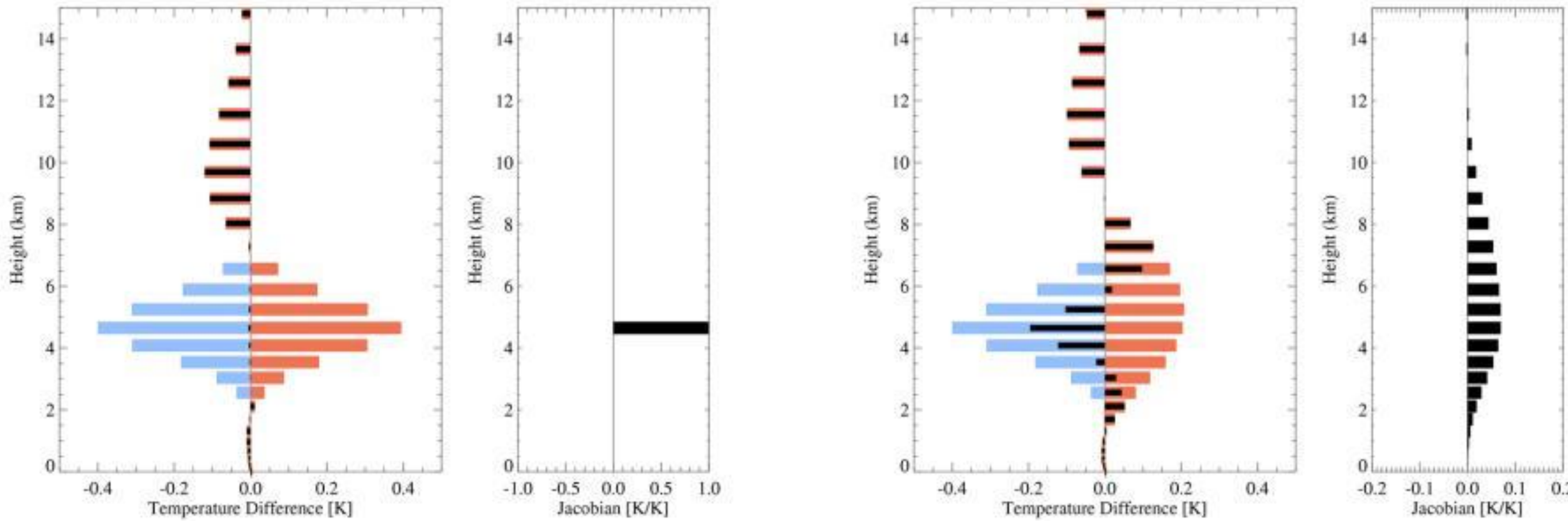
- Comparison with **radiosondes** (best estimate of truth but limited coverage)
- Comparison of e.g. 48hr and 24hr forecasts (so called ***NMC method***)
- Comparison of **ensembles** of analyses made using perturbed observations
- **Flow-dependent** "error of the day" routine estimation

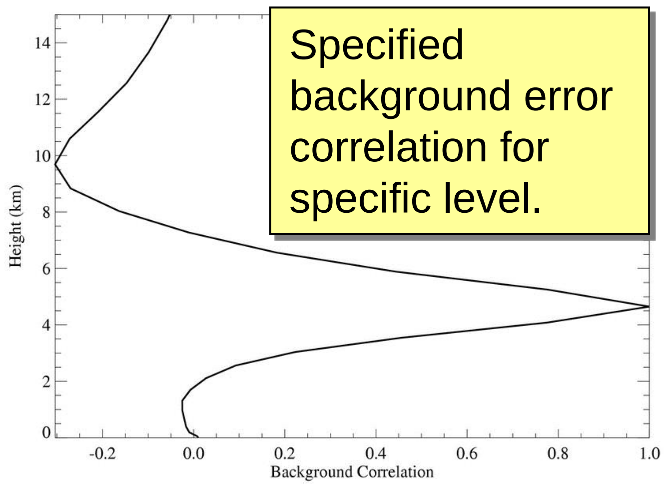
Temperature background error correlations with 700 hPa level:



Broad errors
in the mid-lat

Example of background constraint





Specified background error correlation for specific level.

Error in background

Increments

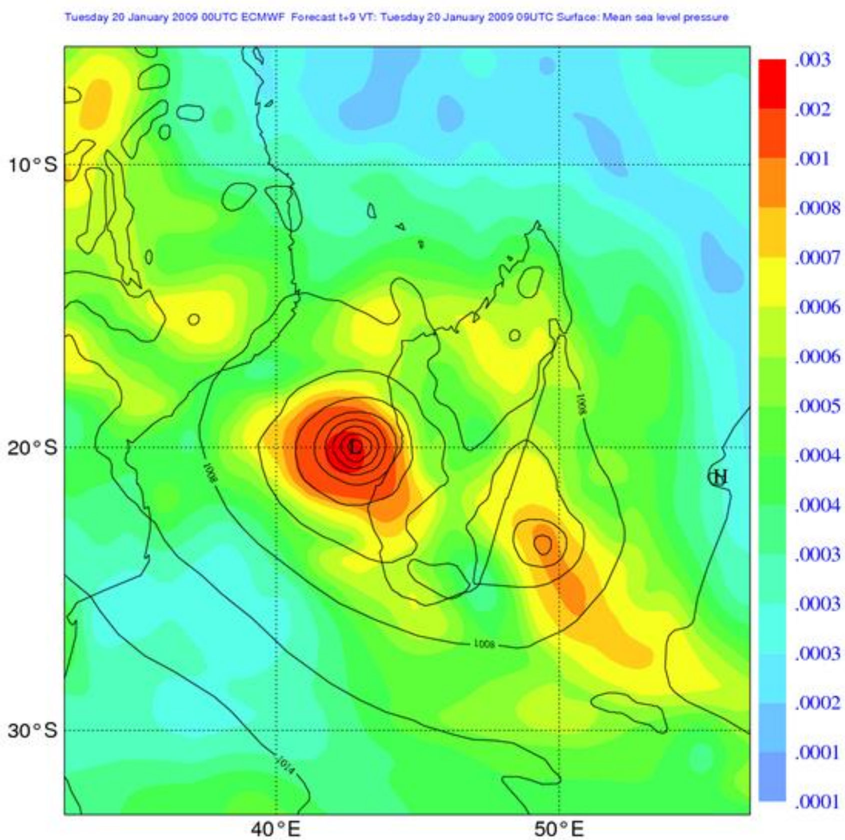
Error in analysis

ECMWF

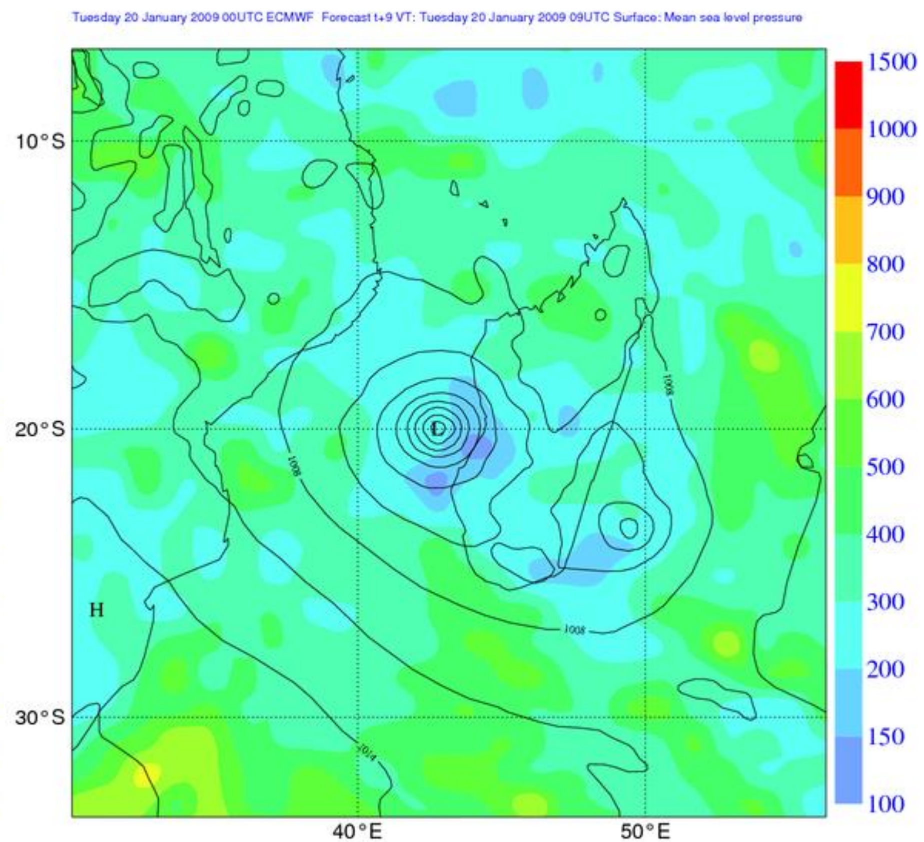
Ensemble of Data Assimilations

- Use of correlation information from the EDA in 4D-Var

EDA StDev of LNSP



EDA Lscale of BG errors LNSP



ECMWF

3.) Systematic errors and bias correction

Systematic errors (biases)

Systematic errors (or biases) must be removed before the assimilation otherwise biases will propagate in to the analysis (causing **global damage** in the case of satellites!).

$$\boxed{\textit{Bias} = \textit{mean}\left[\, Y_{obs} - H(X_b)\,\right]}$$

Observed radiance

RT model

Background atmospheric state

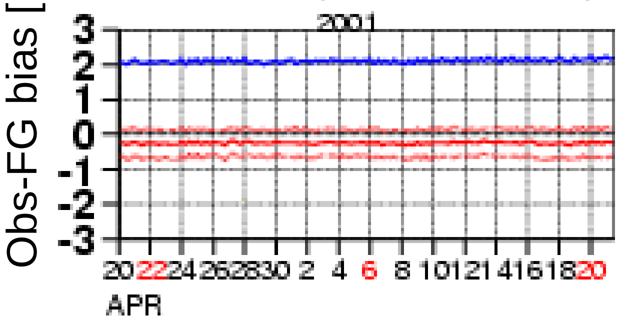
Sources of systematic error in radiance assimilation include:

- Instrument error (calibration)
- Radiative transfer error (spectroscopy or RT model)
- Cloud/rain/aerosol screening errors
- Systematic errors in the background state from the NWP model

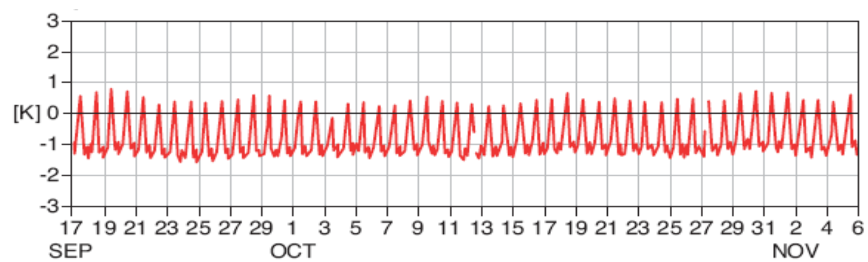
What kind of biases do we see? (I)

Biases are obtained from long-term monitoring of observation minus background.

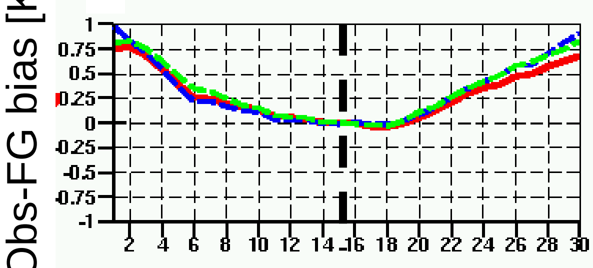
Constant bias (HIRS channel 5)



Diurnal bias variation in a geostationary satellite



Bias depending on scan position (AMSU-A ch 7)

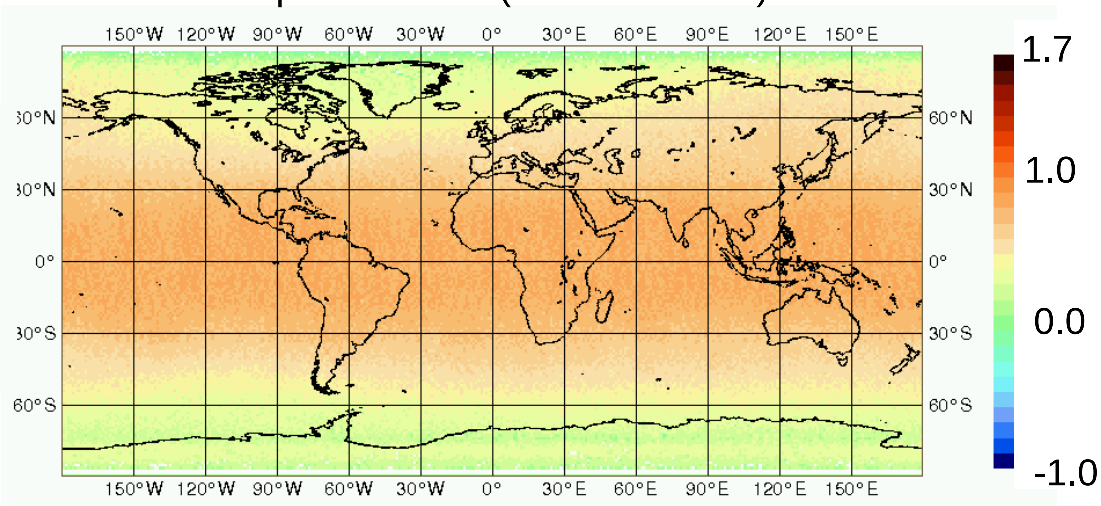


high zenith angle

nadir

high zenith angle

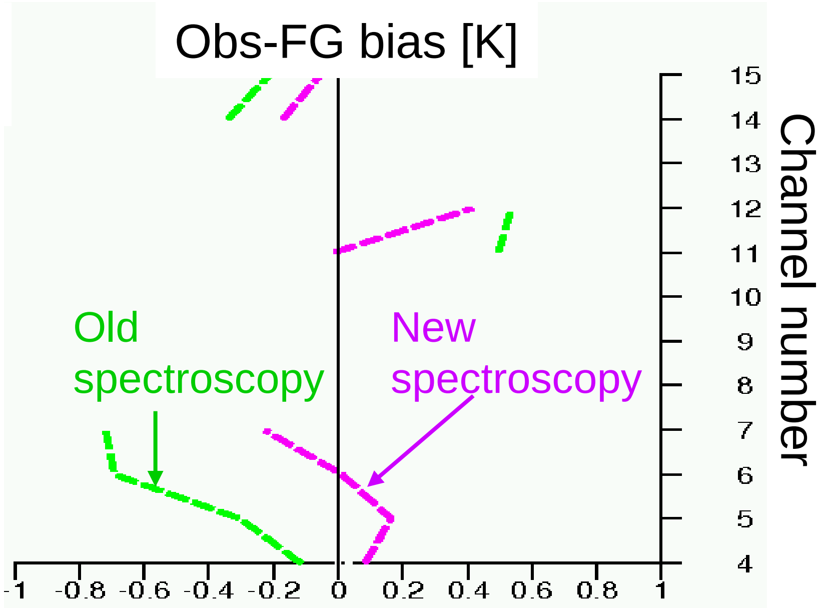
Air-mass dependent bias (AMSU-A ch 10)



ECMWF

What kind of biases do we see? (II)

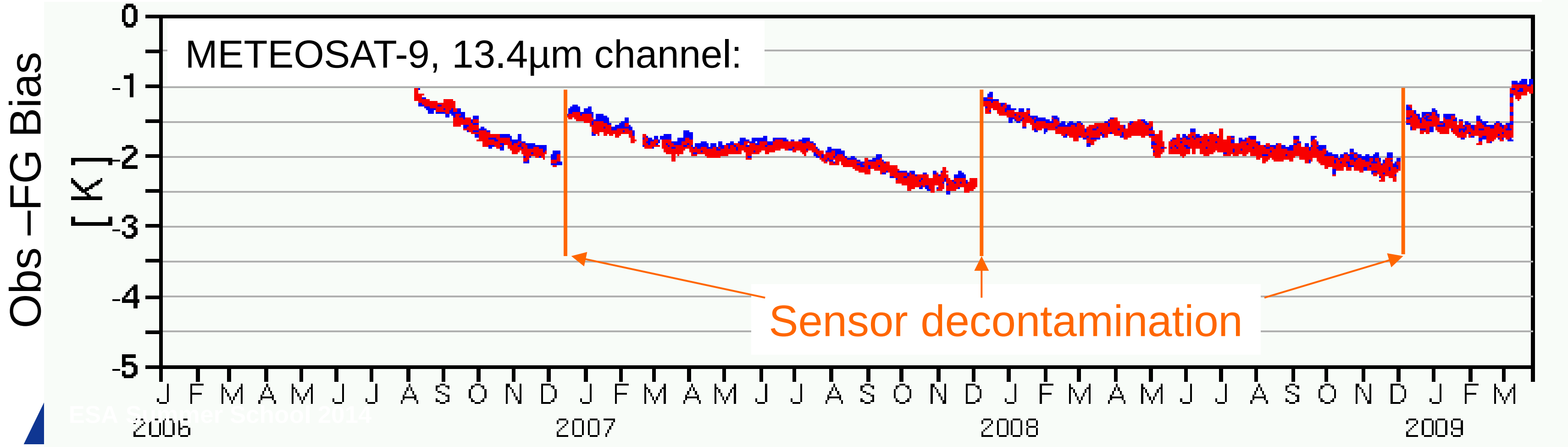
Different bias for HIRS due to different spectroscopy in the radiative transfer model:



Other common causes for biases in radiative transfer:

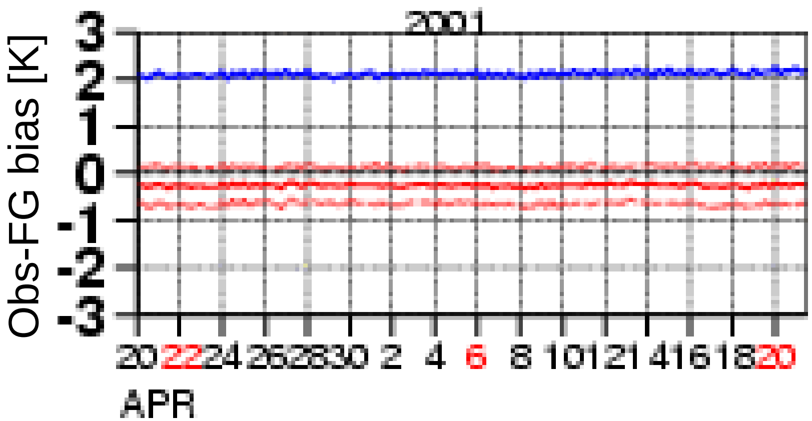
- Bias in assumed concentrations of atmospheric gases (e.g., CO_2)
- Neglected effects (e.g., clouds, aerosols)
- Incorrect spectral response function
- ….

Drift in bias due to ice-build up on sensor:

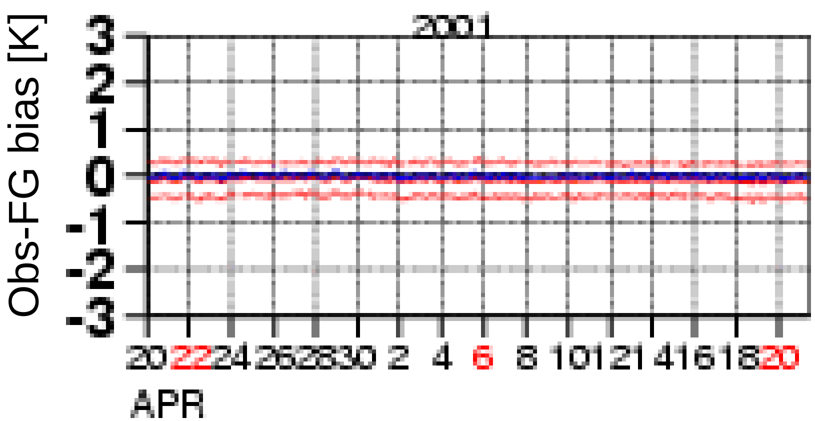


Diagnosing the source of bias (I)

Monitoring the background departures (averaged in time and/or space):



HIRS channel 5 (peaking around 600hPa) on NOAA-14 satellite has +2.0K radiance bias against FG.



Same channel on NOAA-16 satellite has no radiance bias against FG.

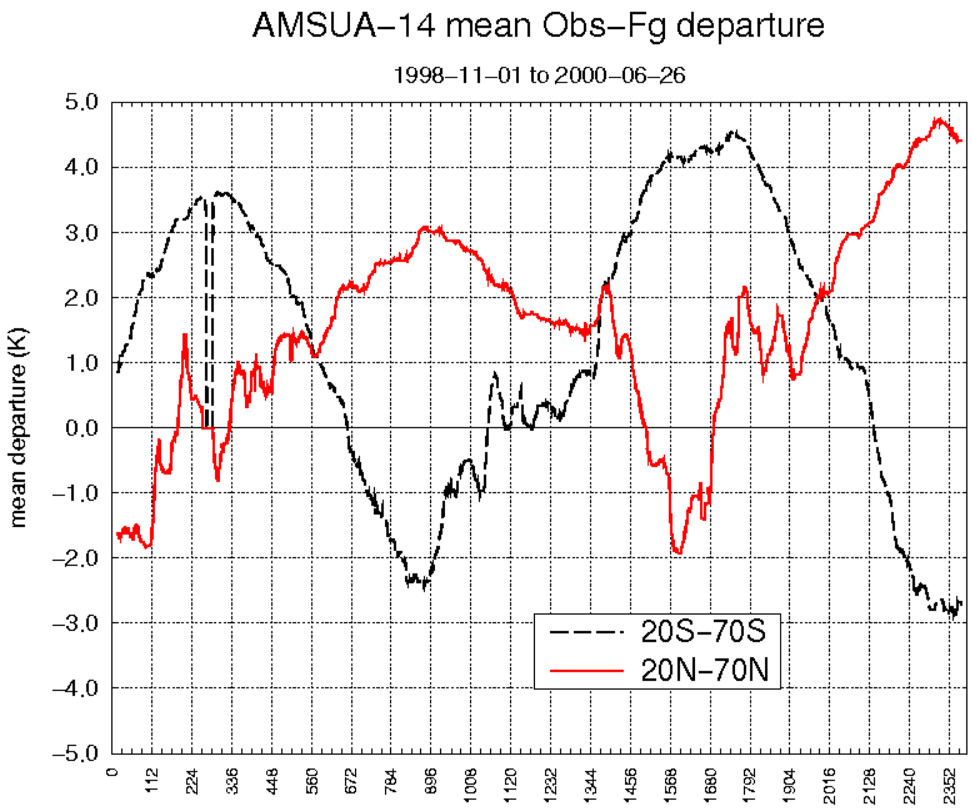
→ **NOAA-14 channel 5 has an instrument bias.**

Diagnosing the source of bias (II)

What about biases in the forecast model?

This time series shows an **apparent time-varying bias** in AMSU channel14 (peaking at 1hPa).

By checking against other research data (HALOE and LIDAR data) the bias was confirmed as an **NWP model temperature bias** and the channel was assimilated with **no bias correction**

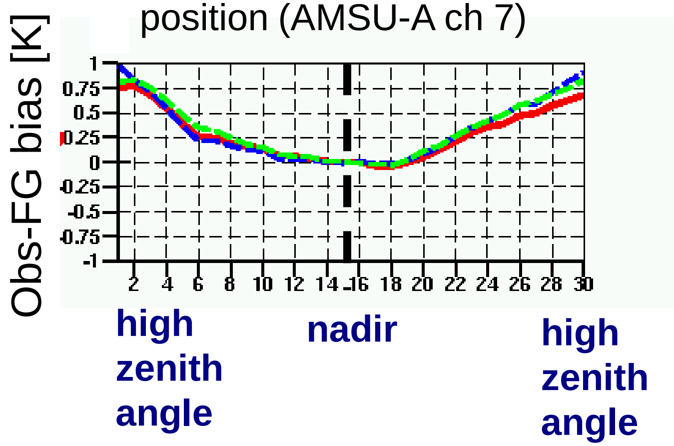


ECMWF

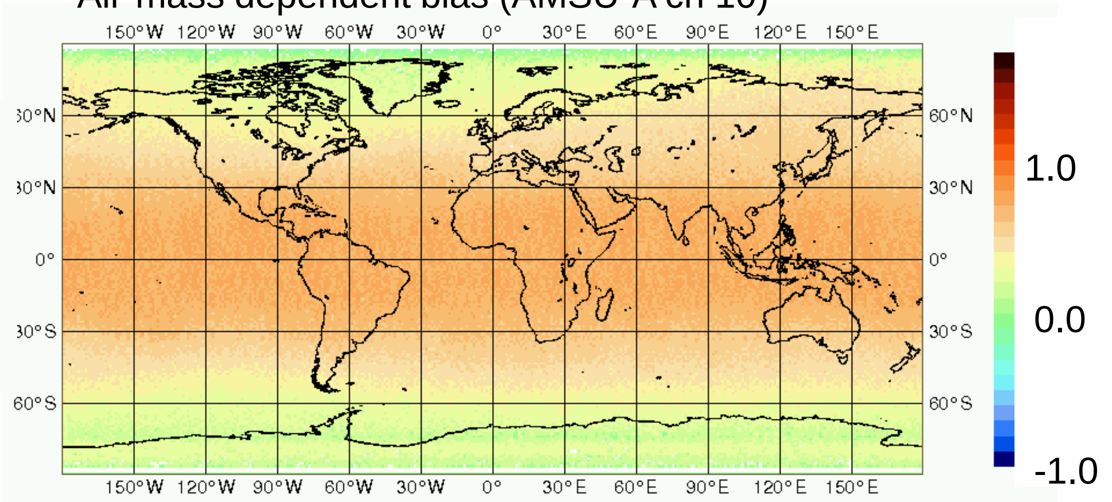
Bias correction

- Biases need to be corrected before or during the assimilation.
- Usually based on a "model" for the bias, depending on a few parameters.
  - ➢ Ideally, the bias model "corrects only what we want to correct".
  - ➢ If possible, the bias model is guided by the physical origins of the bias.
  - ➢ Usually, bias models are derived empirically from observation monitoring.
- Bias parameters can be estimated offline or as part of the assimilation ("variational bias correction")

Bias depending on scan position (AMSU-A ch 7)

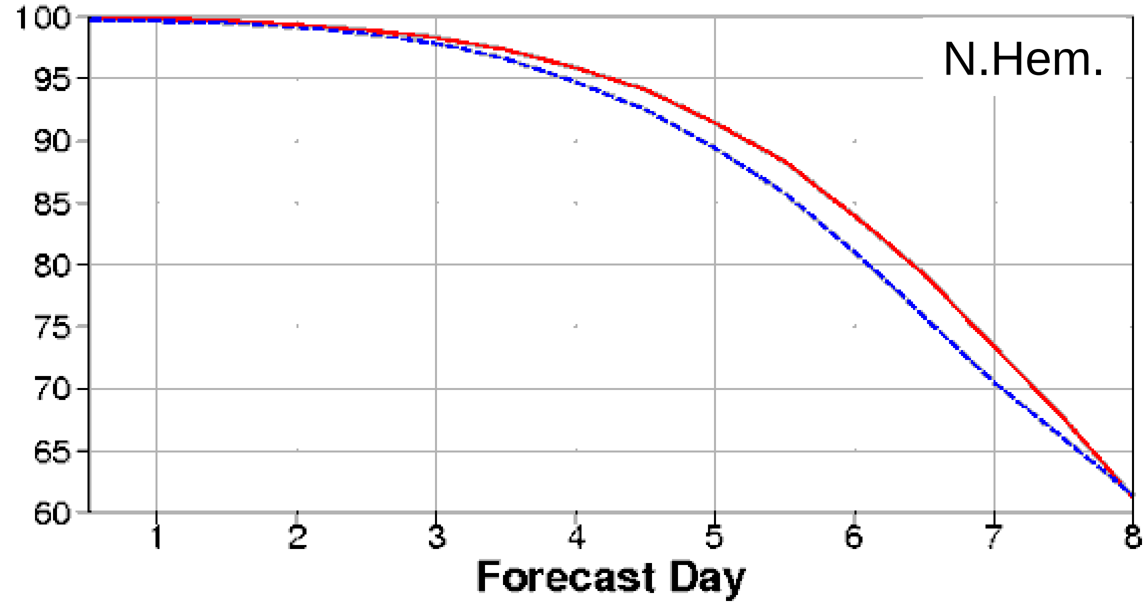


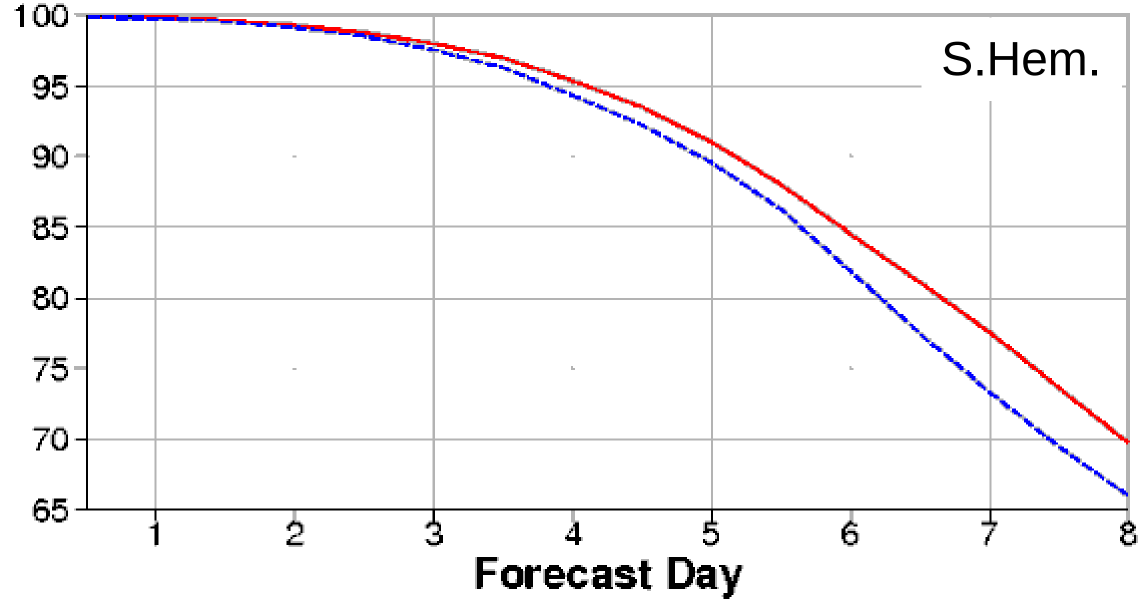
Air-mass dependent bias (AMSU-A ch 10)



Importance of bias correction

Forecast impact comparing
operational bias correction
vs
bias correction with static global constant only





4.) Ambiguity in radiance observations

Ambiguity between geophysical variables

When the primary absorber in a sounding channel is a **well mixed gas** (e.g. oxygen) the radiance essentially gives information about variations in the **atmospheric temperature profile only**.

$$L(\nu) = \int_0^\infty B(\nu, T(z)) \left[\frac{d\tau(\nu)}{dz} \right] dz$$

When the primary absorber is **not well mixed** (e.g. water vapour, ozone) the radiance gives **ambiguous information** about the temperature profile and the absorber distribution. This ambiguity must be resolved by:

- Differential channel sensitivity
- Synergistic use of well mixed channels (constraining the temperature)
- The background error covariance (+ physical constraints)

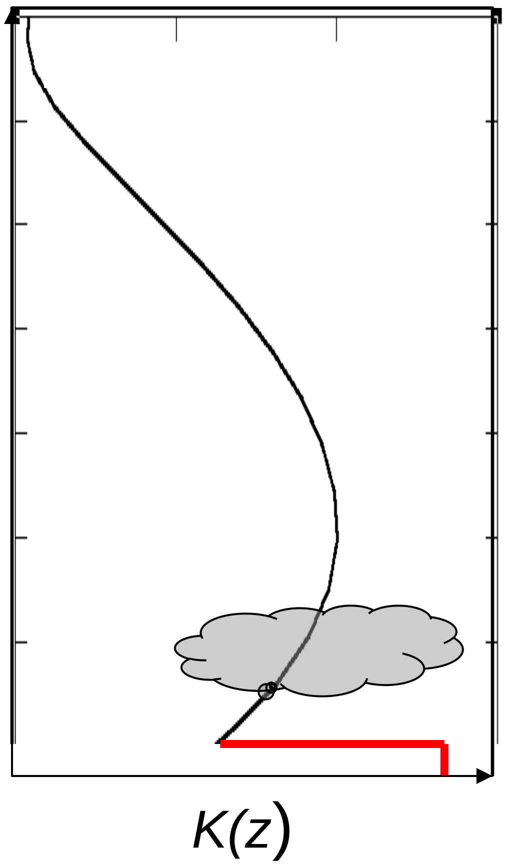
Ambiguity with surface and clouds

By placing sounding channels in parts of the spectrum where the absorption is **weak** we obtain temperature (and humidity) information from the **lower troposphere** (low peaking weighting functions).

BUT …

These channels (obviously) become more sensitive to surface emission and the effects of cloud and precipitation.

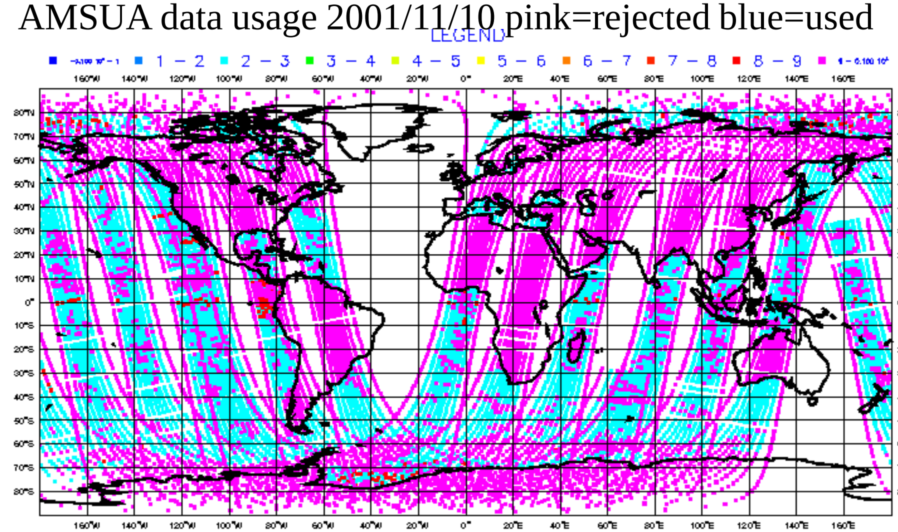
In most cases **surface or cloud** contributions will **dominate the atmospheric signal** in these channels and it is difficult to use the radiance data **safely** (i.e. we may alias a cloud signal as a temperature adjustment).



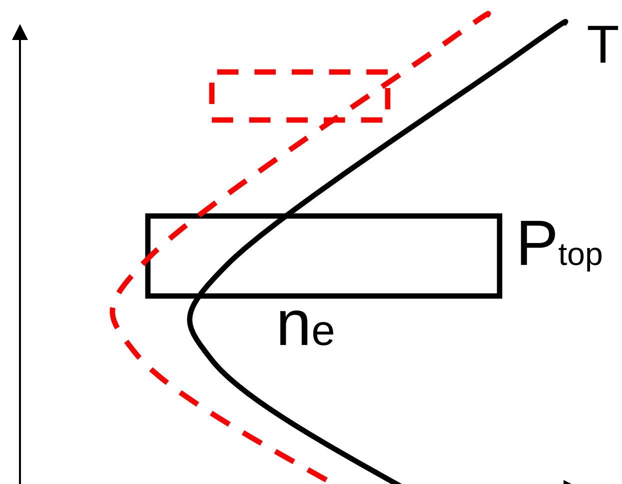
Options for using lower-tropospheric sounding channels

• Screen the data carefully and only use situations for which the surface and cloud radiance contributions can be computed very accurately *a priori* (e.g. cloud free situations over sea). But meteorologically important areas are often cloudy!

AMSUA data usage 2001/11/10 pink=rejected blue=used



• Simultaneously estimate atmospheric temperature, surface temperature / emissivity and cloud parameters within the analysis or retrieval process (need very good background statistics !). Can be dangerous.



5) Quality Control

Data selection and quality control (QC):

The primary purpose of this is to ensure that the observations entering the analysis are consistent with the assumptions in the observations error covariance (R) and the observation operator (H).

Primary examples include the following:

Rejecting bad data with gross error (not described by R)

Rejecting data affected by clouds if H is a clear sky RT
Thinning data if no correlation is assumed (in R)

Always blacklisting data where we do not trust our QC!

Data selection and quality control (QC):

Often checks are performed using the forecast background as a reference. That is an observations is rejected if the departure from the background exceeds a threshold T_{QC}:

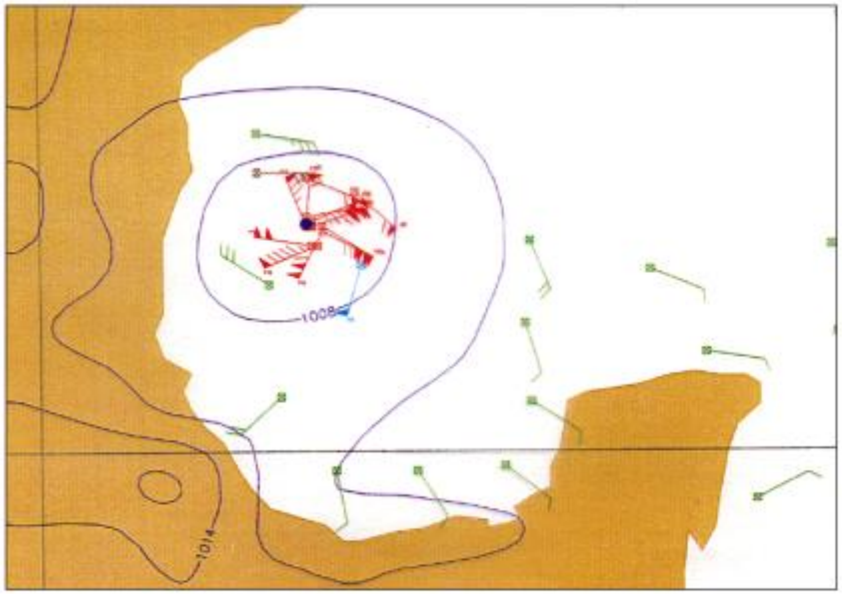
$$Y_{obs} - H(X_b) > T_{QC}$$

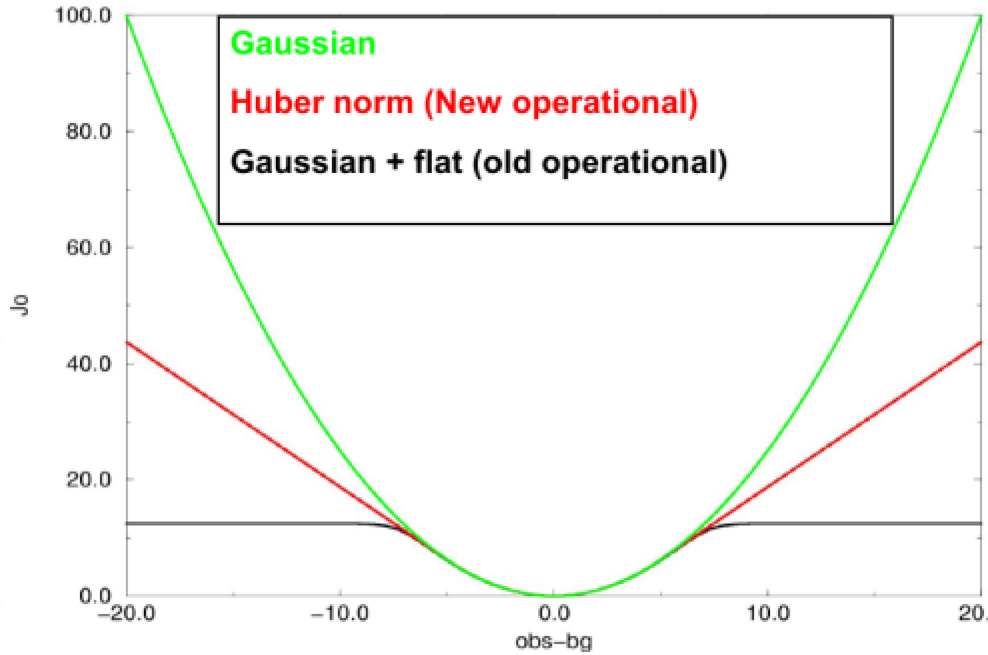
But sometimes large errors in the background can lead to:

- False rejection of a good observation
- Missed rejection of a bad observation

Data selection and quality control:

- False rejection of a **good** observation



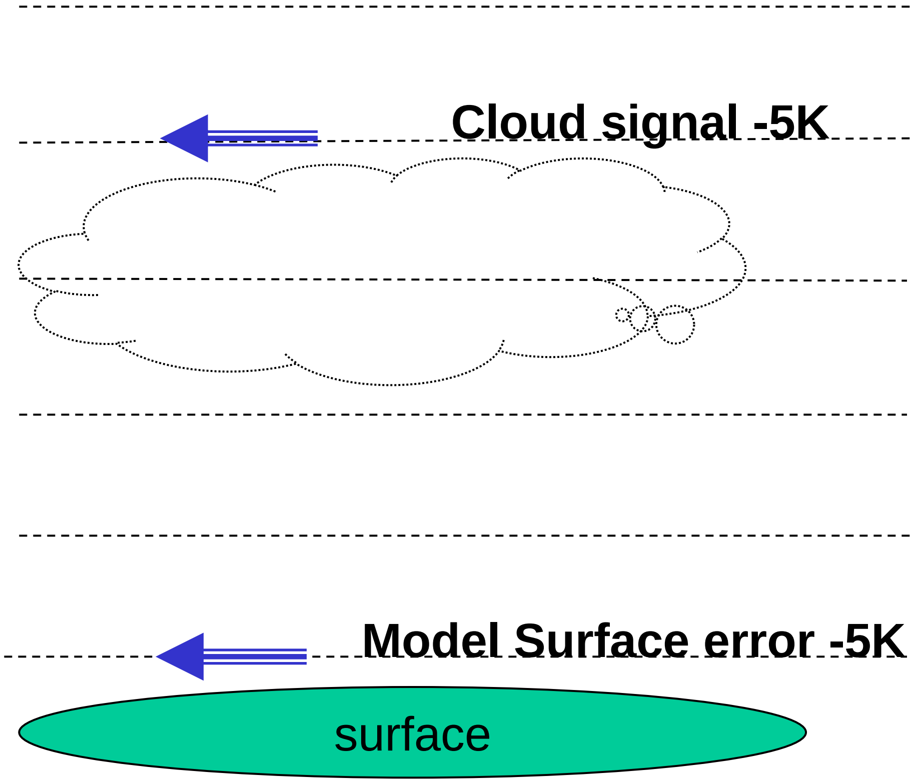


The **numerous** failing observations are good, but a bad background is causing them to be rejected. We **need** these observations to improve the analysis !

Instead of rejecting, we give the observations a lower weight so **collectively** they can influence and improve the analysis. In this framework a single bad observation would do no damage.

Data selection and quality control:

- Missed rejection of a **bad** observation



The radiance are contaminated by cloud (**cold 5K**) compared to the clear sky value.

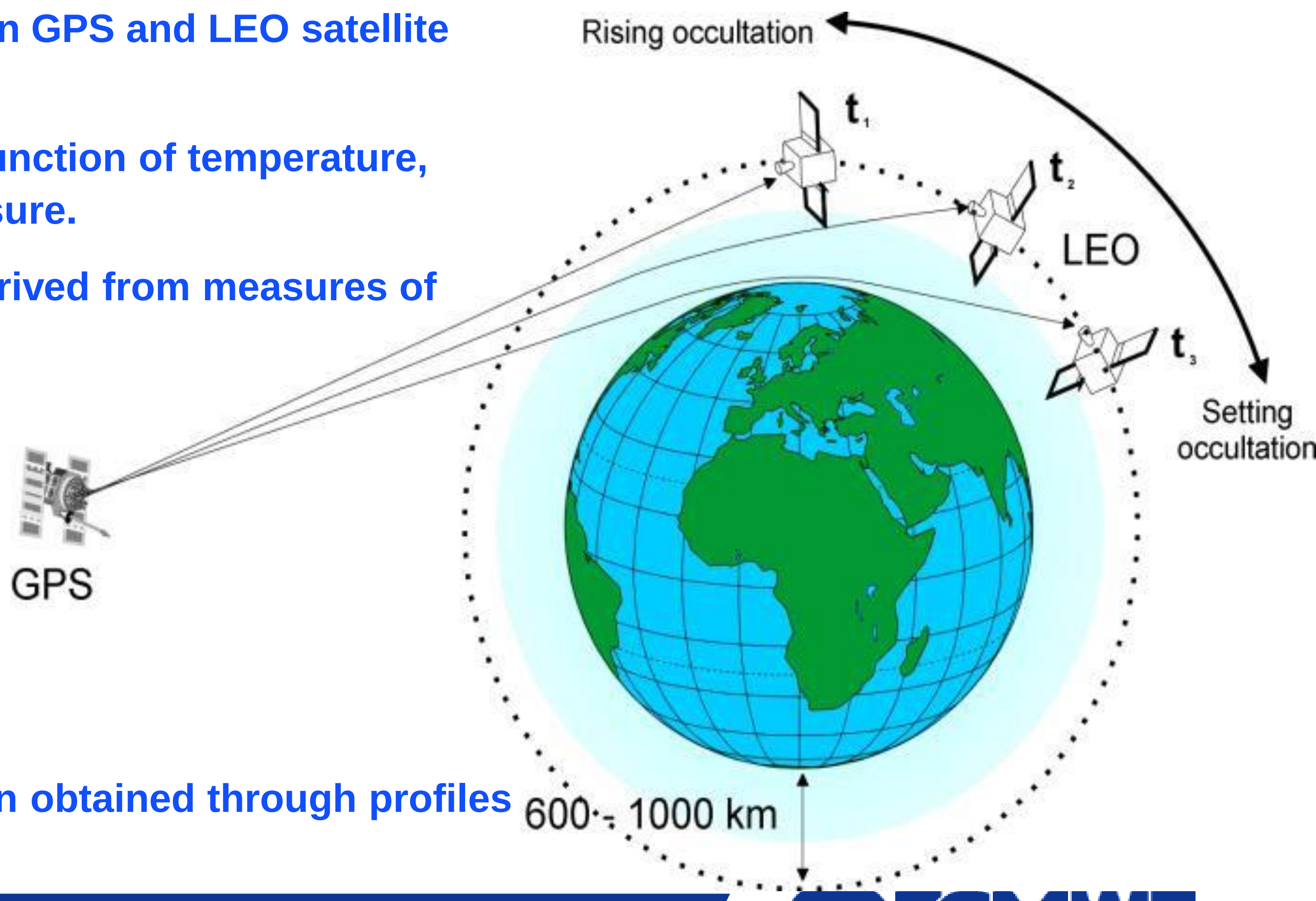
But our computation of the clear sky value from the background is also **cold by 5K** due to an error in the surface skin temperature.

Thus our checking (against the background) sees no reason to reject the observation and is it **passed**!

6.) Some other observation types

GPS (GNSS) Radio occultation

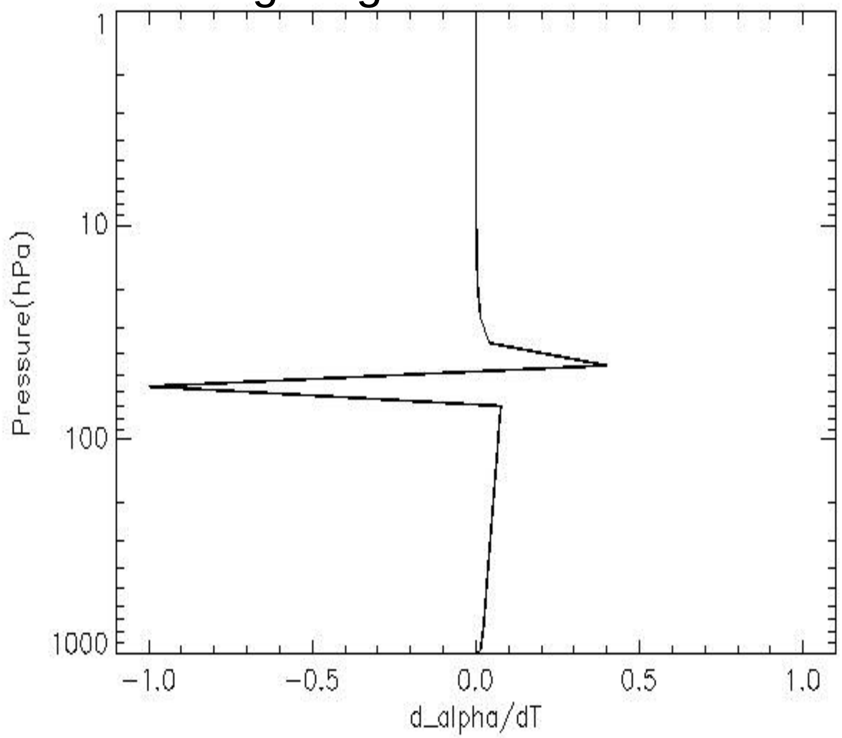
- Limb measurement.
- Gradients in refractivity cause bending of a signal path between GPS and LEO satellite (Snell' s law).
- Refractivity is a function of temperature, humidity and pressure.
- Bending angle derived from measures of phase delay.



- Profile information obtained through profiles of bending angles.

GPS RO characteristics

1d weighting function:



- All weather-capability:
    - Not affected by cloud or rain.

- Largely bias-free. Can help "anchor" bias corrections for radiances.

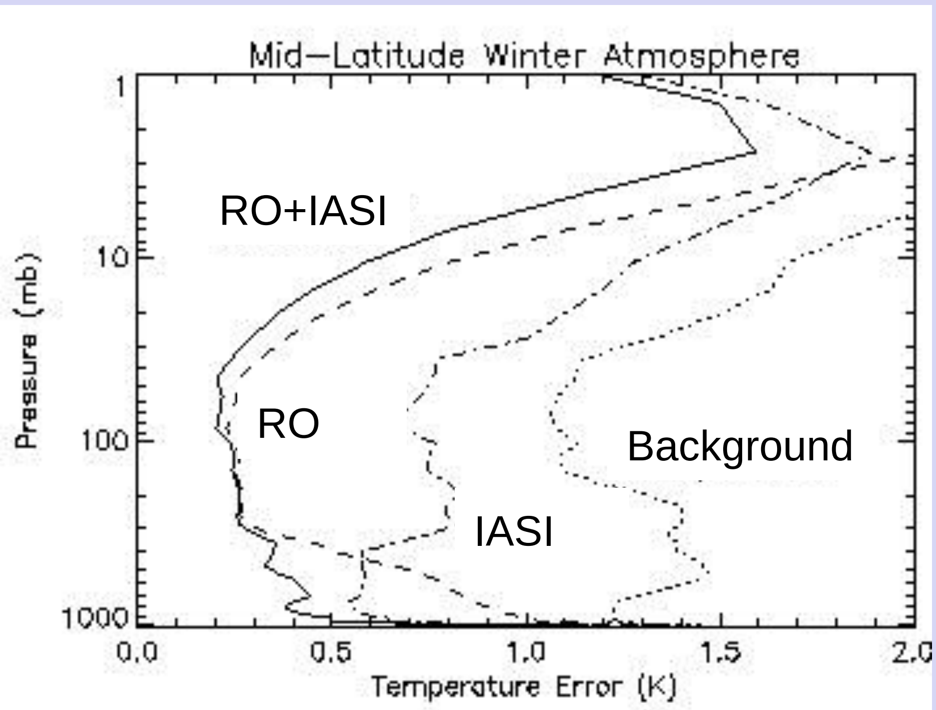
- Good vertical resolution. Can see error structures that nadir radiances can't.

- But has broad horizontal weighting function! - Around 70% of the bending occurs over a ~450km section of ray-path, centred on the tangent point *(point closest to surface).*

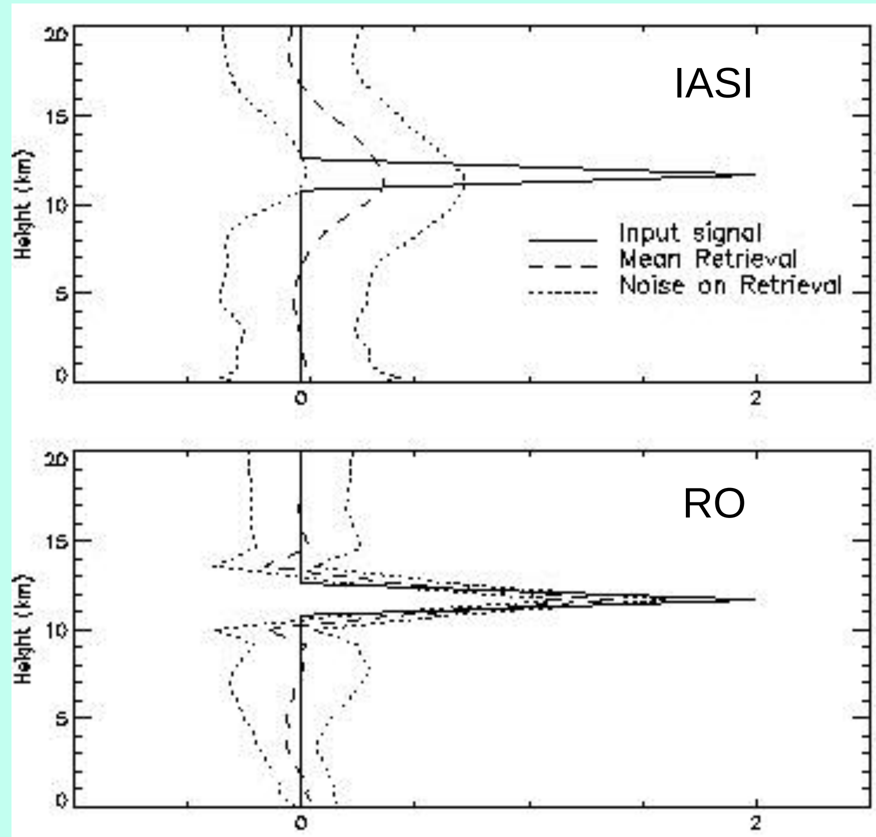
GPS RO vs IASI: 1DVAR simulations

See Healy and Collard 2003, QJRMS:

Expected retrieval error:

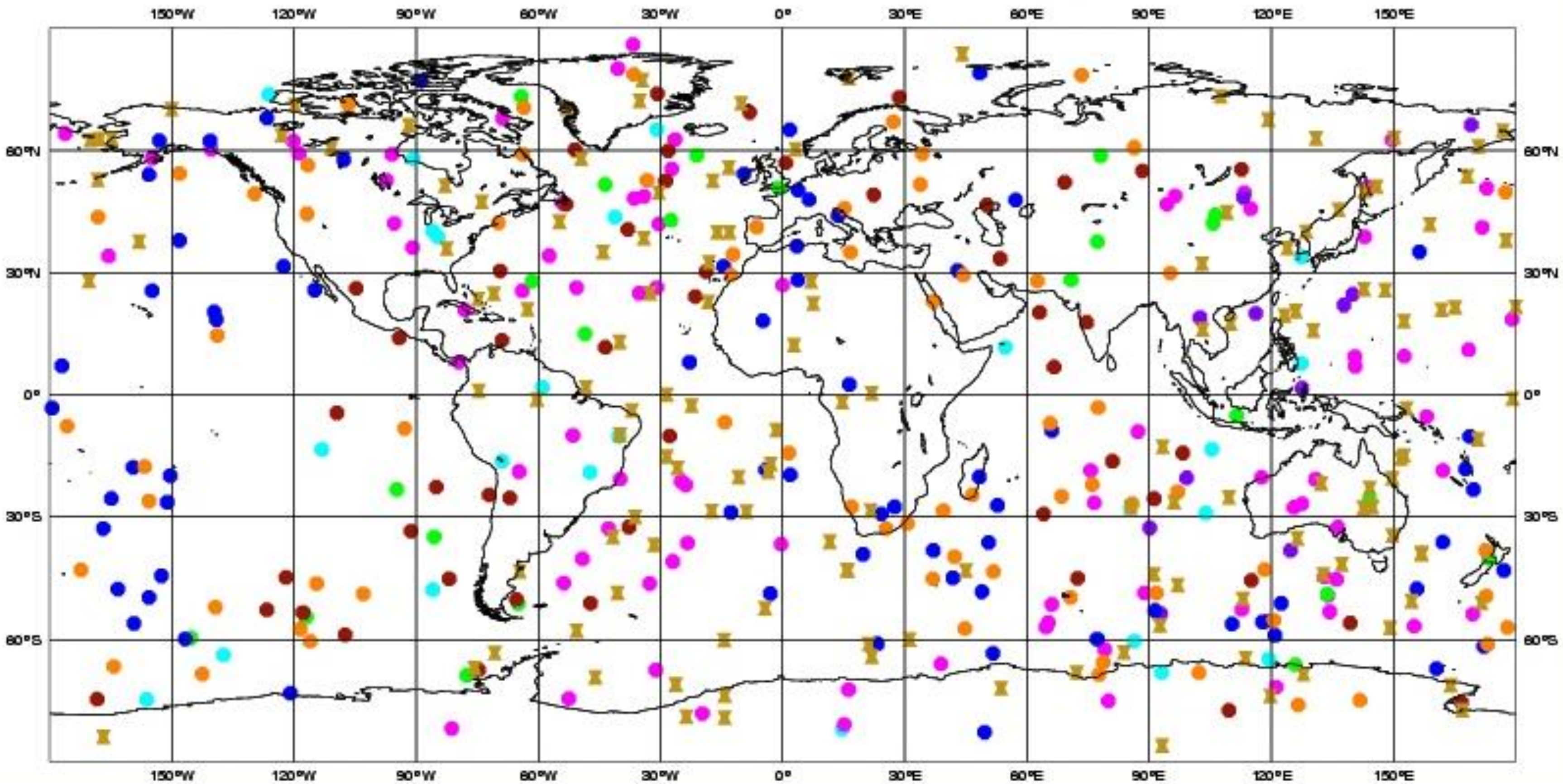


Power to resolve a peak-shaped error in background:

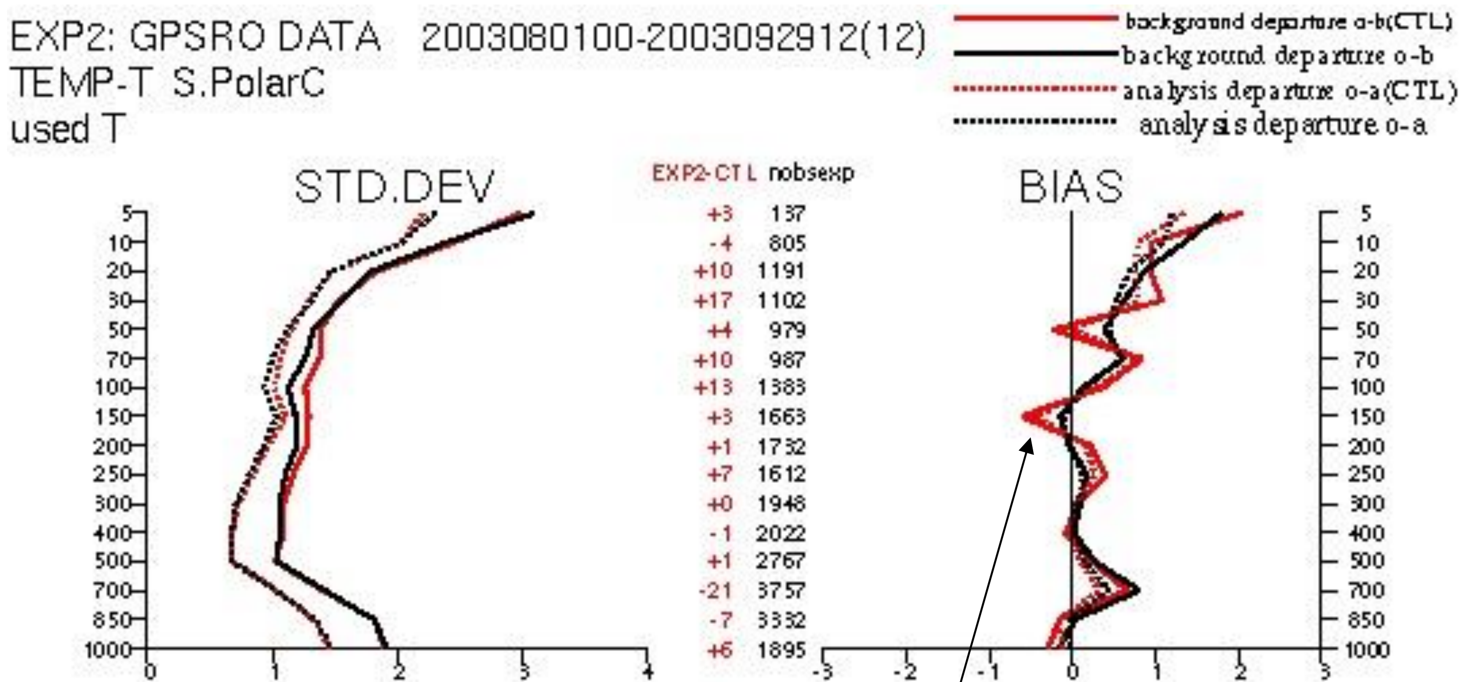


GPS RO data coverage in 6-hour period

Data from GRACE-A, GRAS,
COSMIC-1, COSMIC-2, COSMIC-3, COSMIC-4, COSMIC-5, COSMIC-6



Radiosonde comparisons for Antarctica 12h forecasts



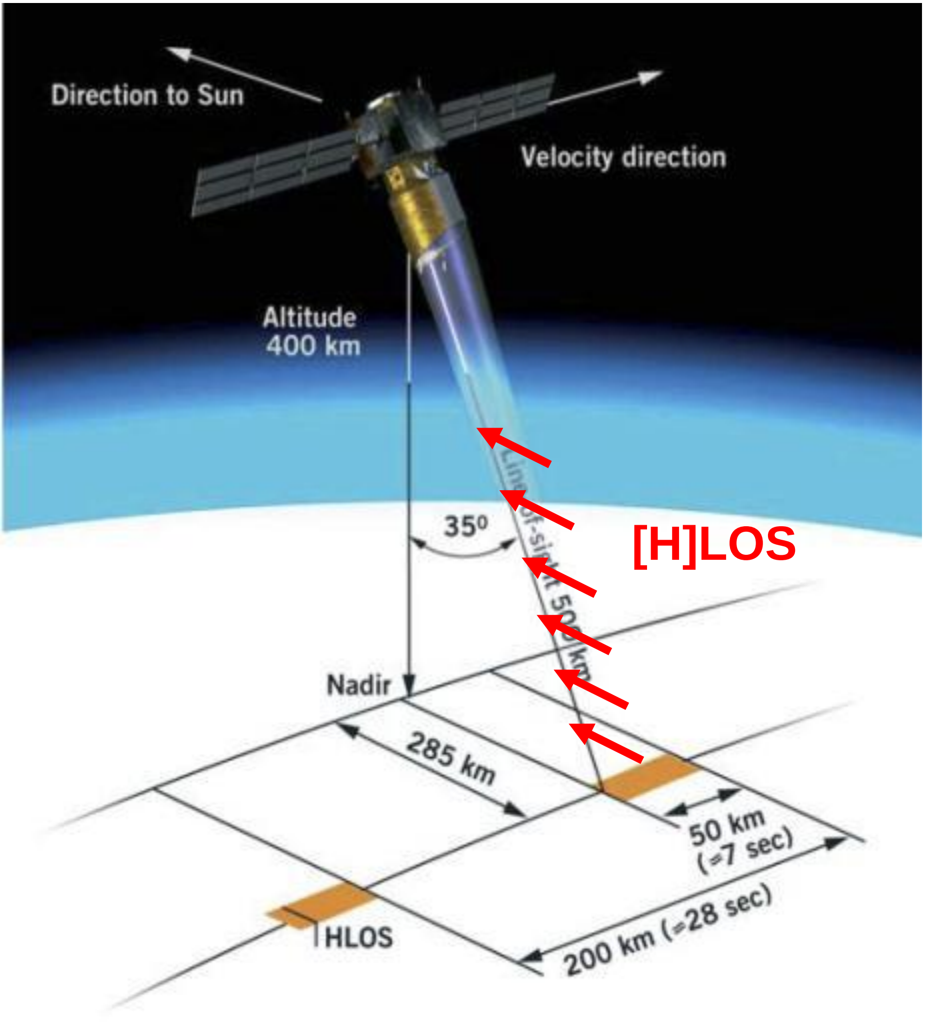
Red lines: Without GPSRO

Black: With GPSRO

Structure in the mean fit thought to be caused by inconsistencies in the AIRS and AMSU bias corrections schemes

ECMWF

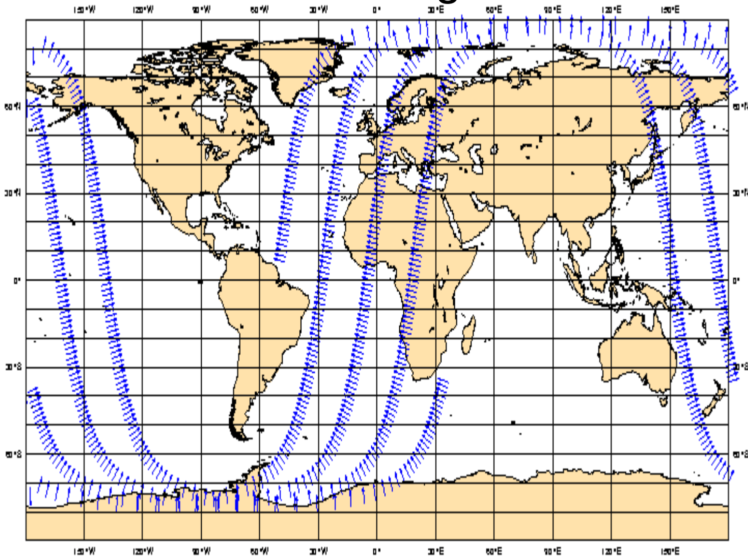
Atmospheric Dynamics Mission ADM-Aeolus

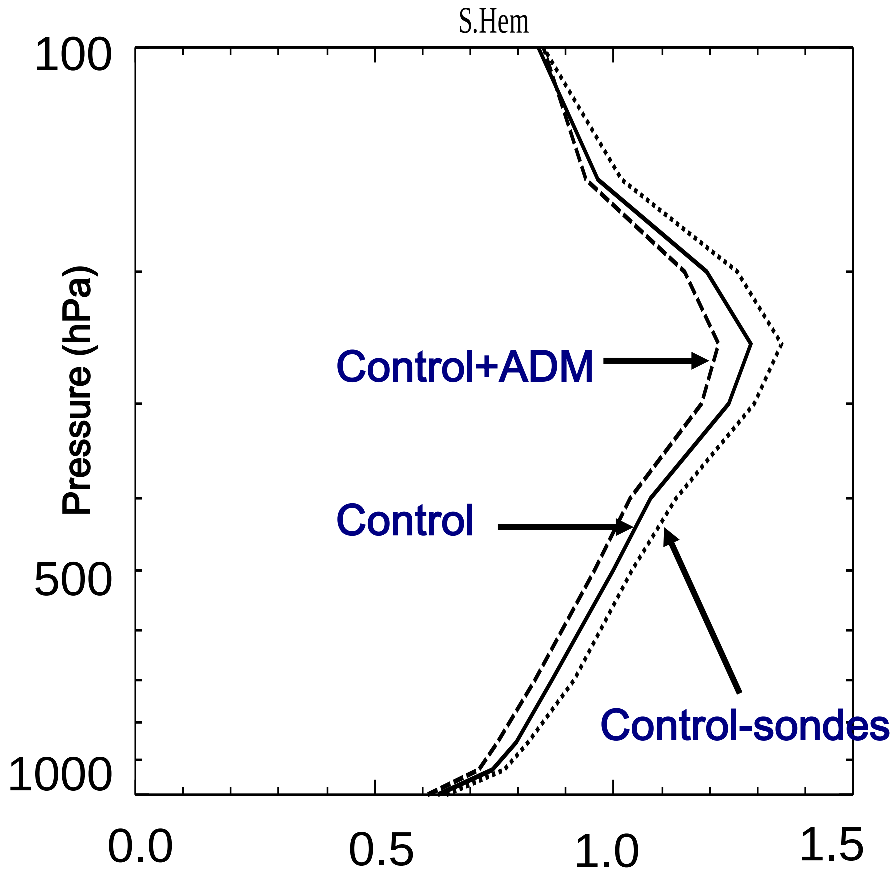


- **ESA Earth Explorer**
- **Doppler wind lidar** to measure line-of-sight (LOS) wind profiles in the troposphere to lower stratosphere (up to 30 km)
- **Vertical resolution** from 250 m - 2 km
- **Horizontal averages** over 50 km every 200 km
- Requirements on **random error** of horizontal LOS wind:
  - <1 m/s (z=0-2 km, for Δz=0.5 km)
  - <2 m/s (z=2-16 km, for Δz= 1 km
- **First wind lidar in space**; will also give information on aerosol/cloud optical properties.
- Launch: not before end of 2015

ADM-Aeolus: Simulated impact

6-hour data coverage:





Zonal wind forecast error (m/s) for 12-hour forecast.

Expected forecast impact for ADM-Aeolus has been simulated using ensemble methods.

Simulated DWL data adds value at all altitudes and well into longer-range forecasts.

Summary

The assimilation of satellite radiance observations has a very powerful impact upon NWP data assimilation schemes, but…

- **BACKGROUND ERROR STRUCTURES**
(what are they and are they correctly specified?)

- **SYSTEMATIC ERRORS**
(what are they and are they correctly specified?)

- **AMBIGUITY BETWEEN VARIABLES, QC**
(both atmospheric and surface / cloud contamination)

We must know the error characteristics and limitations of the data to exploit the observations effectively. (H and R).