The Use of Principal Component Analysis for the Assimilation of High-Resolution Infrared Sounder Observations for Numerical Weather Prediction

A.D. Collard

ECMWF
Shinfield Park, Reading, UK
Andrew.Collard@ecmwf.int

ABSTRACT

Methodologies for the efficient representation of observations for high-resolution infrared sounders for the purposes of assimilation into numerical weather prediction models are discussed. The use of principal component analysis is explored and it is noted that while the available information in the observations is stored efficiently the non-locality of the Jacobians that arise may cause practical problems in an operational assimilation system. Reconstructed radiances appear to be a more realistic approach in the near term. However, initial experiments with reconstructed radiances do not appear to give significant improvements in forecast skill above that already demonstrated with the current use of advanced sounder data.

1 Introduction

In the last few years, kilochannel infrared sounders have become a major part of the assimilation systems of many numerical weather prediction centres (see, for example, the data observation impact studies of Kelly and Thépaut , 2007).

Currently two such instruments are operationally assimilated at NWP centres. The first was the Atmospheric InfraRed Sounder (AIRS) which was launched in May 2002 on the EOS-AM (Aqua) satellite. EOS-AM is a research satellite rather than an operational meteorological one but the data have been distributed in near-real-time to NWP centres allowing the observations to be used operationally. This has provided NWP centres with four years experience with this data type before the first data from the true operational missions arrived (e.g., McNally *et al.* (2006); Collard *et al.* (2003); Auligné *et al.* (2003); Le Marshall *et al.* (2006)). AIRS measures the radiance emitted from the Earth in 2378 channels covering the spectral interval from 650–2665cm^{-1} at a resolution of around 1cm^{-1}. It is unique amongst this type of sounder as it is a grating spectrometer rather than an interferometer.

The Infrared Atmospheric Sounding Interferometer, IASI, (Chalon *et al.*, 2001) is an infrared Fourier transform spectrometer which was first launched on the MetOp-A satellite in October 2007. It is the first truly operational kilochannel infrared sounder. It has 8461 channels covering the spectral interval from 645–2760cm^{-1} at a resolution of 0.5cm^{-1} (apodised). IASI radiances are currently being assimilated operationally at a number of NWP centres (e.g., Collard and McNally (2009); Hilton *et al.* (2009a); Hilton *et al.* (2009b)).

The CrIS instrument will be the equivalent instrument to IASI on the NPOESS satellite series. It will measure 1400 channels in the 635-2450cm^{-1} range with a spectral resolution between 1.125 and 4.5cm^{-1}. It is scheduled to be launched in January 2011.

The high volume of data resulting from these observations presents many challenges, particularly in the areas of data transmission, data storage and assimilation. In all cases, these observations are currently assimilated at numerical weather prediction centres using around 300 channels or fewer.

As there are fewer pieces of independent information in a IASI spectrum than channels by around two orders of magnitude (e.g., Huang *et al.*, 1992), there is potential to manipulate these observations such that the bulk of the information is retained while reducing the data volume significantly. Various methods have been proposed to do this including the use of principal component analysis which is the subject of this paper.

The next two sections outline the theoretical properties of the principal components of the noise-normalised radiances and their derived product of reconstructed radiances. Section 4 discusses the practical use of reconstructed radiances in an operational numerical weather prediction system.

2 Principal Component Analysis

Efficiently representing the relatively small amount of information contained in thousands of IASI channels is particularly suited to Principal Component Analysis (PCA). This technique (Huang and Antonelli, 2001; Antonelli *et al.*, 2004; Aires *et al.*, 2002a) relies on the use of a dataset of thousands of spectra representing the full range of meteorological conditions from which the the principal components are calculated.

The principal components are generally calculated from a representative set of spectra which have been noise-normalised and from which the mean of the population has been subtracted. The noise normalisation ensures that noise is not being fit selectively in spectral regions with low signal-to-noise ratios. It is also normal to use radiances rather than brightness temperatures as the instrument noise is more constant in radiance space (in addition where signal-to-noise ratios are very high, such as the short wavelength part on the IASI spectrum, negative radiances can occur which cannot be represented as brightness temperatures).

If this set of radiances is represented by the $n_{ch} \times n_{obs}$ matrix $\mathbf{X}$ (where n_{ch} is the number of channel and n_{obs} the number of observations and hence each column of $\mathbf{X}$ represents an individual spectrum), the covariance, $\mathbf{C}$ of these spectra is given by

$$\mathbf{C} = \frac{1}{n_{obs}}\mathbf{X}\mathbf{X}^T = \mathbf{L}\mathbf{\Lambda}\mathbf{L}^T \tag{1}$$

where $\mathbf{L}$ is the matrix of principal components and $\mathbf{\Lambda}$ their associated eigenvectors.

Alternatively, $\mathbf{X}$ can be decomposed through thin singular value decomposition thus:

$$\mathbf{X} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T \tag{2}$$

where $\mathbf{U}$ has dimensions $n_{ch} \times q$; $\mathbf{\Sigma}$, $q \times q$; and $\mathbf{V}$, $n_{obs} \times q$ where q is the smaller of n_{ch} and n_{obs}

As (noting that $\mathbf{V}^T\mathbf{V} = \mathbf{I}$ and $\mathbf{\Sigma} = \mathbf{\Sigma}^T$)

$$n_{obs} \times \mathbf{L}_q\mathbf{\Lambda}_q\mathbf{L}_q^T = \mathbf{X}\mathbf{X}^T = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T\mathbf{V}\mathbf{\Sigma}\mathbf{U}^T = \mathbf{U}\mathbf{\Sigma}\mathbf{\Sigma}\mathbf{U}^T \tag{3}$$

$\mathbf{U}$ and $\mathbf{L}_q$ can be seen to be equivalent as can $\mathbf{\Lambda}_q$ and (apart from the n_{obs} scaling factor) $\mathbf{\Sigma}\mathbf{\Sigma}$, where $\mathbf{L}_q$ and $\mathbf{\Lambda}_q$ correspond to the first q eigenvectors/eigenvalues of $\mathbf{L}$ and $\mathbf{\Lambda}$.

Whether one prefers to calculate principal components through the covariance or singular value decomposition therefore depends on the specifics of the problem being addressed. It is desirable to have as many spectra as possible in the training dataset and, when a global training set is needed, typically tens of thousands are used. In this case the $\mathbf{X}$ matrix may get very large and it may be more efficient, particularly in terms of computer memory, to calculate (and optionally update) the covariance matrix $\mathbf{X}\mathbf{X}^T$. The calculations presented in this paper were thus all performed via this route.

The principal component amplitudes $\mathbf{p}$ are related to the observed radiances $\mathbf{y}$ (which are noise normalised observed radiances with the mean observation subtracted) through

$$\mathbf{p} = \mathbf{L}^\mathbf{T}\mathbf{y}. \tag{4}$$

It is generally found that of the order of 200 globally-generated principal components are required to represent the atmospheric signal contained in the original spectrum. The remaining components are mostly noise. The exact number depends upon the application and the tolerance to information loss. This is partially addressed by Turner *et al.* (2006).

If the observation vector has an error ε_y, the corresponding error for the principal component amplitudes is given by

$$\varepsilon_{\mathbf{p}} = \mathbf{L}^{\mathbf{T}} \varepsilon_y \tag{5}$$

leading to an instrument error covariance in principal component space given by:

$$Cov(\mathbf{p}) = E[\varepsilon_{\mathbf{p}} \varepsilon_{\mathbf{p}}^T] = \mathbf{L}^T \mathbf{E} \mathbf{L} \tag{6}$$

where **E** is the error covariance of the original observations. As we are using noise-normalised observations, **E** should be the identity matrix, resulting in

$$Cov(\mathbf{p}) = \mathbf{L}^T \mathbf{L} = \mathbf{I} \tag{7}$$

In practice the noise matrix used for normalisation may differ from the true matrix resulting in $Cov(\mathbf{p})$ deviating from the identity matrix, **I**, and in general having off-diagonal elements (terms in the error budget unrelated to instrument noise such as forward model error will,in general, also have off-diagonal terms).

It should be noted that, as for the reconstructed radiances discussion to follow, the error from the forward model is not propagated through Eqn. 7 and should be a term *additional* to $Cov(\mathbf{p})$.

A number of different types of training set may be used to obtain the principal components. The main decisions to be made are whether to use observed or calculated radiances; a local or global training set; and whether to allow for correlated noise in the noise-normalising process.

Principal component analysis can be performed using simulated or real radiances. The advantage of simulated data is that it can be used when there are no real observations and therefore can be used in pre-flight testing of systems. One also has the flexibility with simulated observations to ensure that the signature of particular atmospheric structures is present in the training set, especially if the simulated data are noise free (i.e., instrument noise is not artificially added).

Simulated data, however, may not contain the same information as reality, as the simulations are limited by the accuracy of the radiative transfer model and how realistic the range of atmospheres used in the training set are. This is particularly true where cloudy radiances are being used. Therefore, if simulated data are used, it may still necessary to either add additional eigenvalues based on real data or to add real spectra into the training set. For this reason, the use of real data is usually preferred, at least for the efficient distribution of the observations where fidelity is paramount.

An example of the use of local rather than global training sets is the production of "granule-based" principal components with AIRS data (Antonelli *et al.*, 2004; Tobin *et al.*, 2007). Here a granule is made up of six minutes of AIRS observations. This ensures that the spectra are represented by the amplitudes of principal components that are truly representative of the spectra and that data compression is more efficient and accurate. However, in a near-real time data dissemination system the principal components will have to be calculated every six minutes and, as they will differ between granules, will need to be distributed along with the observations. Due to this added complexity, granule-based principal components have not been used in an NWP environment. While they may be worthy of further study in this context, only global training sets are considered hereinafter.

The use of the correct correlated error covariance matrix when noise-normalising the spectra is particularly worthy of consideration for interferometers where the radiances have been apodised — a process that introduces strong error correlations between neighbouring channels. Lee and Bedford (2004) address this point but choose to deapodise the level-1c spectrum, thus making the error covariance matrix diagonal. The danger with using a diagonal error covariance to noise normalise apodised IASI radiances is that signals with spectral signatures

on the scale of the instrument resolution (and which are thus smoothed in the apodisation process) will not be represented in the principal components.

Direct assimilation of principal components requires a suitably fast and accurate radiative transfer model (i.e., a model that can calculate principal components directly rather than first calculating 8461 individual radiances). A number of such models are just becoming available, promising methods being the PCRTA approach of Liu *et al.*(2005; see also Havemann, 2009) and the RTTOV based model by Matricardi (2009).

Assimilation of principal components will require a number of new strategies for practical implementation. In particular, the practice of excluding channels that are affected by certain levels and/or species in particular situations (e.g., excluding channels that are affected by the surface over land; excluding channels sensitive to the stratosphere; excluding channels affected by cloud) is more difficult when the "channels" being used are composed of the whole spectrum. This is illustrated in Figures 1–3, where the Jacobians for the ten leading components of the IASI spectrum are presented. It can be seen that the PC Jacobians are extremely non-localised and that each has contributions from both temperature and humidity (as well as other trace gas species not shown). By choosing subsets of the full spectrum for principal component generation it is feasible that these Jacobians can be made more localised and the temperature and humidity signals can be separated out to some degree, but one possible way of using these information is to reproduce the spectrum from the principal components as explained in the next section.

It should be noted, however, that the direct use of principal components in retrieval schemes has been demonstrated by Huang and Antonelli, 2001; Smith *et al.*, 2005 and Aires *et al.*, 2002b, but these schemes either model clear columns directly or explicitly model clouds — something that is still at the early stages of development in an NWP context.

A final issue with the use of principal component amplitudes in NWP is monitoring. The somewhat localised nature of the Jacobians for normal spectral observations allows a rough correlation to be made between the channels and atmospheric altitudes and absorbing species. This is invaluable if one is trying to diagnose problems with the assimilation system and to identify whether the fault lies with the model or the observation. As noted above, this is not the case for principal components. This is by no means an unsurmountable problem — a selection of spectral channels could be monitored in parallel with the principal component amplitudes, for example — but does illustrate that the effect of principal components on the system is far less intuitive than normal spectral radiances.

3 Reconstructed Radiances

Reconstructed radiances $\tilde{\mathbf{y}}$ are calculated from the principal component amplitudes via:

$$\tilde{\mathbf{y}} = \mathbf{L}_p \mathbf{p}_p = \mathbf{L}_p \mathbf{L}^{\mathbf{T}}{}_p \mathbf{y} \tag{8}$$

Here the p subscripts indicate that only the first p principal components are being considered; the rest are assumed to contain primarily noise. If we restrict $\tilde{\mathbf{y}}$ to a subset of N_R channels, by replacing the first $\mathbf{L}$ above with $\mathbf{L}_{\mathrm{NR}}$, those channels will contain all of the information present in the N_p principal components provided $\mathbf{L}_{\mathrm{p}}$ has $\geq N_p$ positive singular values. The minimum criterion for this is that $N_R \geq N_p$ and in practice this criterion is usually sufficient.

Indeed it can be demonstrated experimentally that, for example, if one produces reconstructed radiances from the first 200 principal components, the complete reconstructed spectrum can be derived to approaching machine precision from just the first 200 channels and $\mathbf{L}_p \mathbf{L}^{\mathbf{T}}{}_p$. Therefore the reconstructed radiances for the channels around 650cm^{-1} which measure stratospheric temperatures also contain information on the solar contribution to the near-surface channels around 2600cm^{-1}. The amplitudes of these signals are however very small and maybe ignored when, as is usual, treating the reconstructed channels are proxies for real observations.

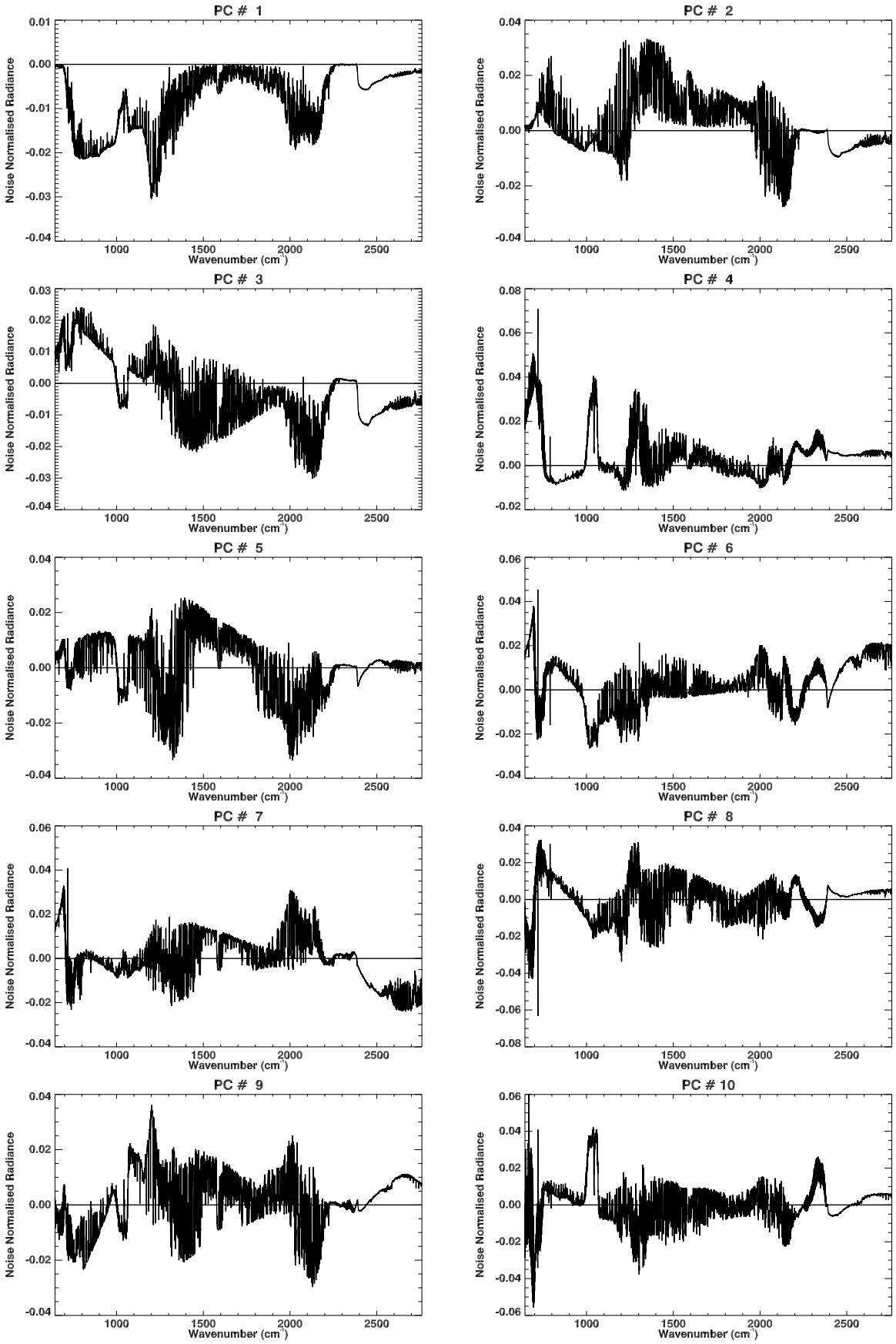


Figure 1: The first ten principal components of the IASI spectrum as generated from 57383 noise-normalised IASI observations from the period between June and November 2007.

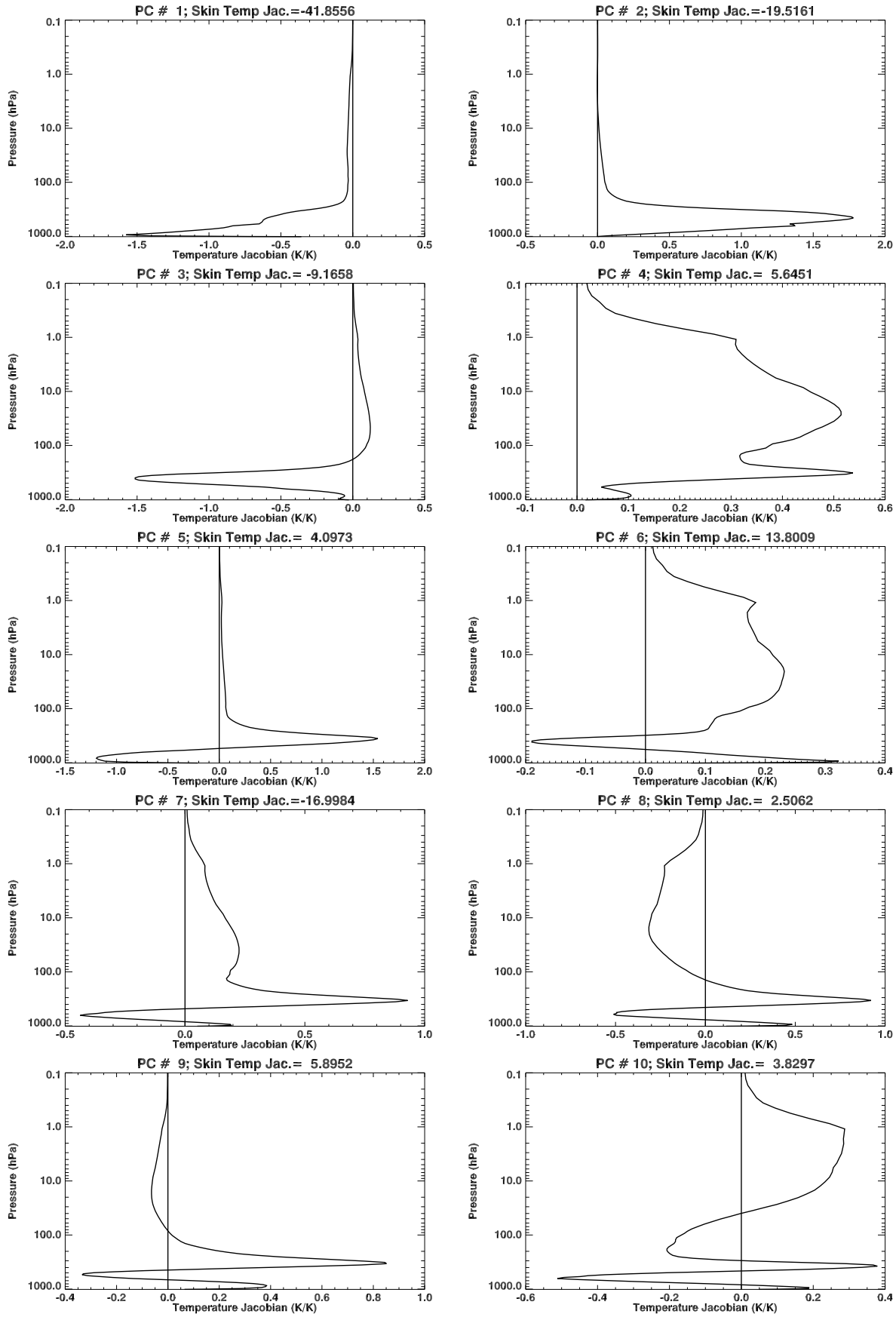


Figure 2: The temperature Jacobians for the first ten principal components of the IASI spectrum shown in Figure 1.

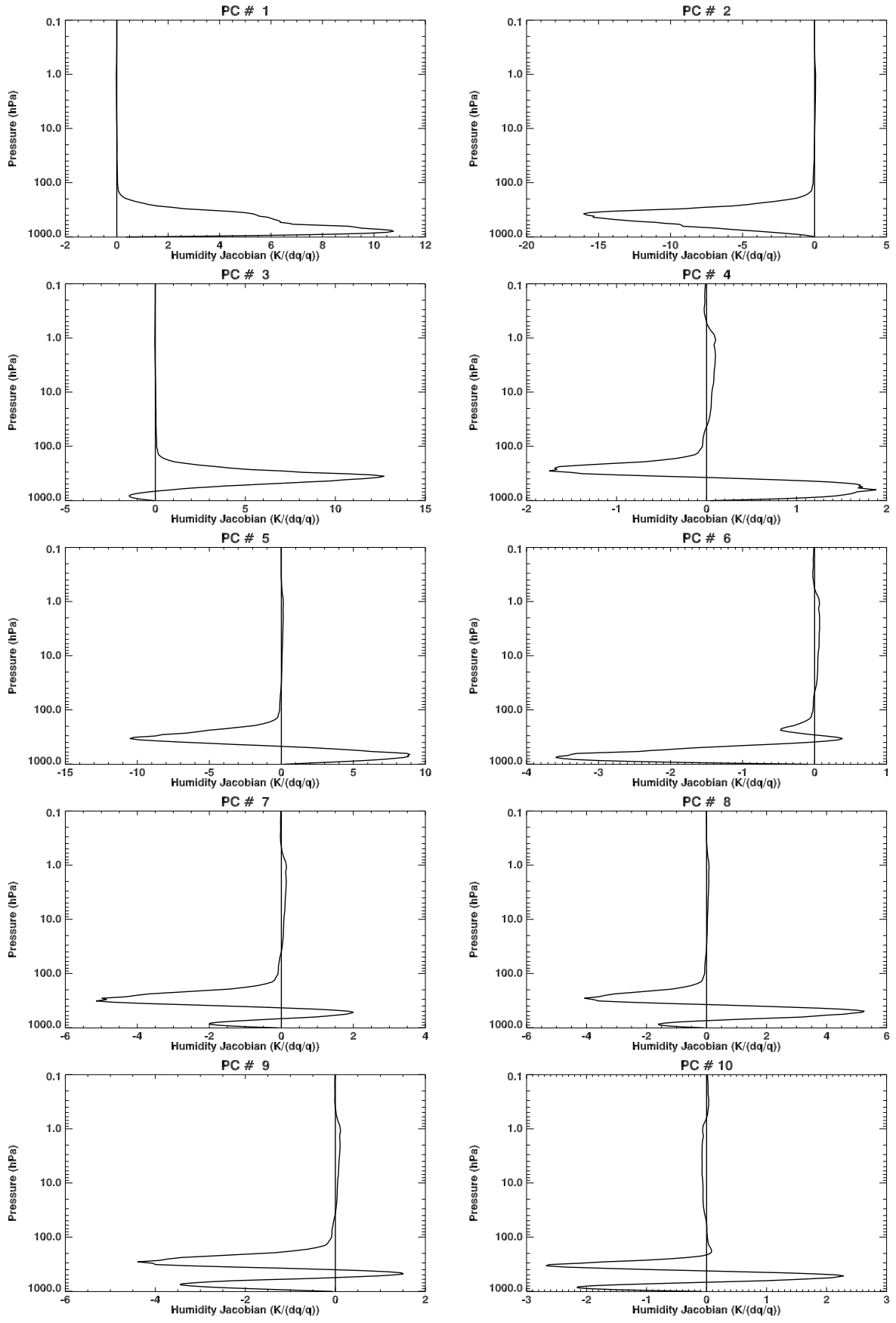


Figure 3: The specific humidity Jacobians for the first ten principal components of the IASI spectrum shown in Figure 1.

One further property of the matrix $\mathbf{L}_p\mathbf{L}^{\mathbf{T}}{}_p$ is that it is *idempotent*. That is $\tilde{\mathbf{y}} = \mathbf{L}_p\mathbf{p}_p = \mathbf{L}_p\mathbf{L}^{\mathbf{T}}{}_p\tilde{\mathbf{y}}$; there is no additional effect on the matrix $\mathbf{y}$ if it is premultiplied by $\mathbf{L}_p\mathbf{L}^{\mathbf{T}}{}_p$ more than once.

In an analogous manner to Eqn. 5 above, the observation error for the reconstructed radiances can be derived from the raw observational error through

$$\varepsilon_{\tilde{\mathbf{y}}} = \mathbf{L}_{\mathrm{NR}}\mathbf{L}^T\varepsilon_{\mathbf{y}} + \varepsilon_{\mathbf{R}} \tag{9}$$

This is the equation when one has a forward model that models the true reconstructed radiances (i.e., the linear combination of all the channels in $\mathbf{y}$). Usually, we wish to treat the reconstructed radiances as estimates of the true radiances but with the noise represented by the unused principal components removed. The additional term, $\varepsilon_{\mathbf{R}}$, is the error introduced on reconstructing the true (error free) spectrum (i.e., this error term comes from the true atmospheric signal that is contained in the unused components). If the forward model that is used explicitly calculates reconstructed radiances, ε_R is zero (although, of course, forward model error remains). If one assumes that the reconstructed radiances are a "noise-filtered" proxy for the original radiances this term will be non-zero but, through careful choice of the number of EOFs and the training set used to produce them, should be at or below the level of the first term in the equation.

Therefore

$$Cov(\tilde{\mathbf{y}}) = E[\varepsilon_{\tilde{\mathbf{y}}}\varepsilon_{\tilde{\mathbf{y}}}^T] = \mathbf{L}\mathbf{L}^T\mathbf{E}\mathbf{L}^T\mathbf{L} + \mathbf{F}_R, \tag{10}$$

where $\mathbf{E}$ is the original observational error covariance matrix (in noise normalised radiance units) and $\mathbf{F}_R$ is the reconstruction error. Again, if the radiances are normalised by the true noise, $\mathbf{E}$=$\mathbf{I}$.

Reconstructed radiances will tend to have channel-correlated errors (even if the original radiance errors are uncorrelated) as the reconstructed channels are derived from the whole spectrum. Figure 4 illustrates how the diagonal of the instrument noise for IASI is greatly reduced on using reconstructed radiances, but Figure 5 illustrates how error correlations between channels are introduced (this figure is for the 366 channel subset that are operationally monitored at ECMWF as this shows the 15μm CO_2 band in greater clarity). That is, the reduction in standard deviation is achieved through combining information from the whole spectrum into each channel, with the result that errors in channels become correlated. For this reason the reconstructed radiances method is best thought of as an *intelligent smoothing* process.

4 Experience with the Assimilation of AIRS Reconstructed Radiances.

AIRS radiances have been assimilated operationally at ECMWF since October 2003 (McNally *et al.*, 2006) using data supplied in near-real time by NOAA/NESDIS. Only 324 channels out of the total of 2378 AIRS channels are supplied, for communication bandwidth reasons. Out of these available channels a maximum of 155 channels are actually assimilated in the configuration used here (CY29R2). The remaining channels are blacklisted for various reasons including sensitivity to solar radiation during the day, and unmodeled non-LTE effects.

The forecast impact from AIRS is significant as illustrated in Figure 6 where the anomaly correlation of 500hPa geopotential height forecast is seen to be much improved in the Southern Hemisphere compared to a system denied AIRS data (the Northern Hemisphere is close to neutral).

The observation noise assumed in the operational assimilation system is plotted in 7 along with the estimated instrument noise for AIRS. The observation noise is assumed to be diagonal and, in places, is significantly higher than the AIRS instrument noise itself. It has been found that significant reductions of this assumed noise can degrade the fit to other observation types and produce worse forecasts. The inflated observation noise reduces the influence of errors arising from unmodeled channel-to-channel correlation in the true error covariance matrix; forward model error (including cloud detection errors); representivity error; and spatial error correlations.

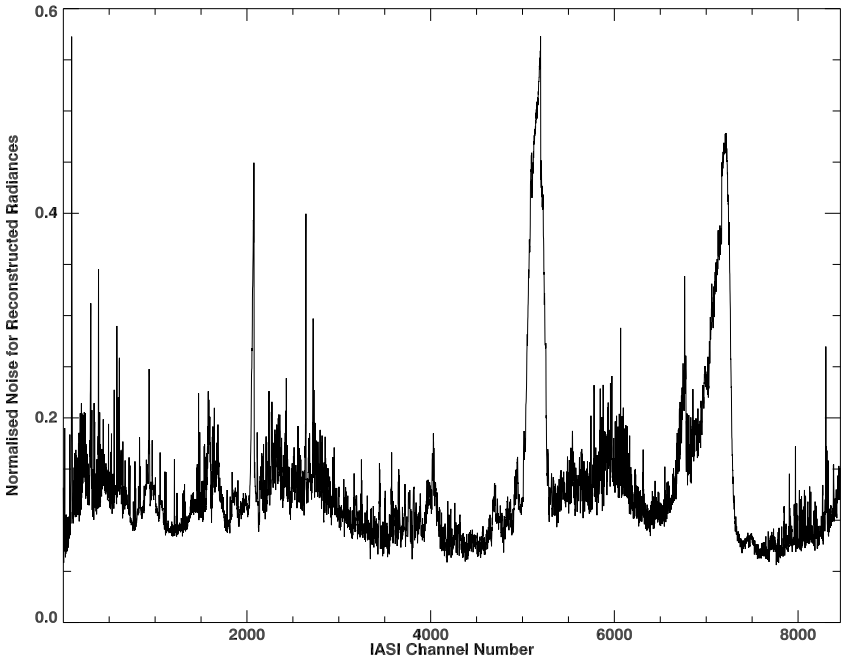


Figure 4: The diagonal of the instrument noise error covariance matrix for reconstructed radiances for a set of 200 principal components derived from real IASI observations (c.f. Eqn 10). "Normalised" here refers to the fact that the original radiances are noise normalised and therefore the original instrument noise covariance is the identity matrix.

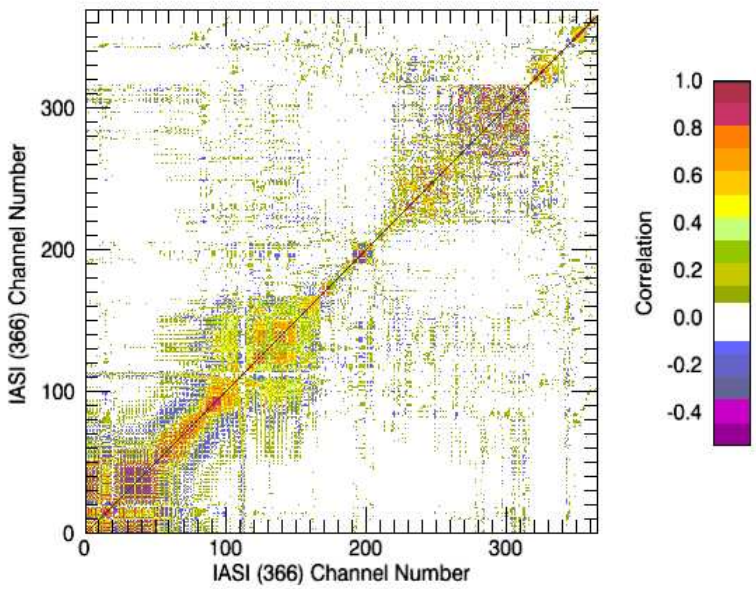
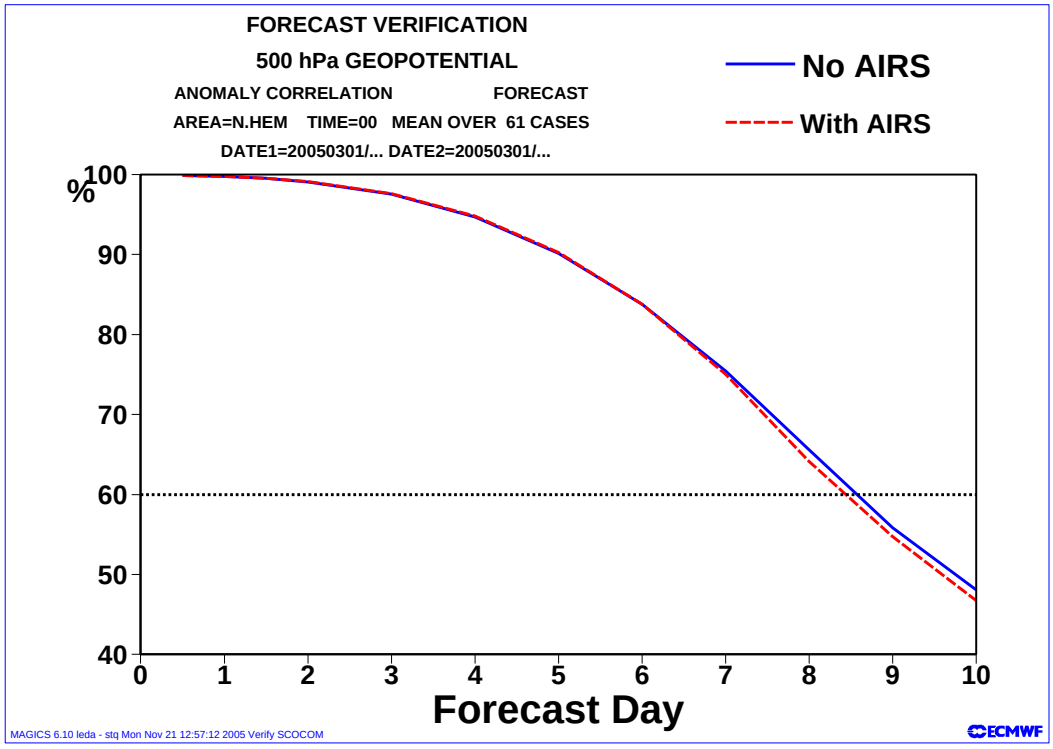


Figure 5: The instrument noise error correlation matrix for the same case as Fig. 4 but only for the 366 IASI channels that are monitored at ECMWF. The approximate spectral bands are as follows: Channels 1–170, 15μm CO_2 band; 171–190 & 208–220, long wave window; 191–207, Ozone; 221–338, water band; 339–358, 4.3μm CO_2 band; 359–366, shortwave window.



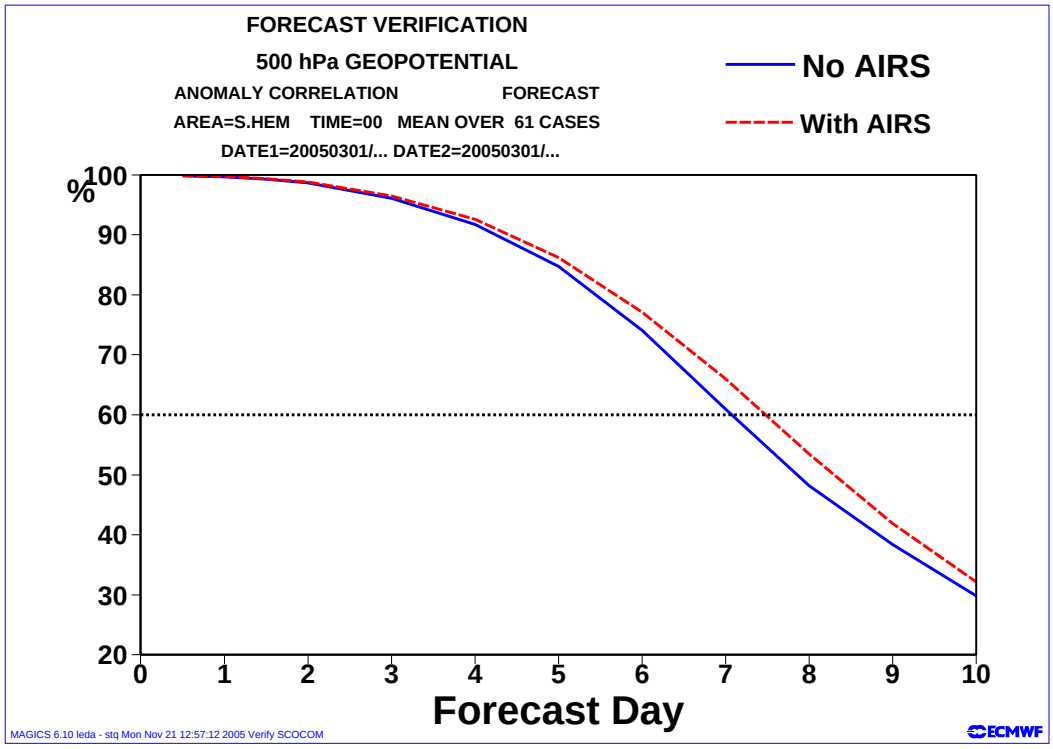


Figure 6: The impact on 500hPa forecast scores on the assimilation of AIRS radiances for the Southern Hemisphere. The impact for the Northern Hemisphere is close to neutral.

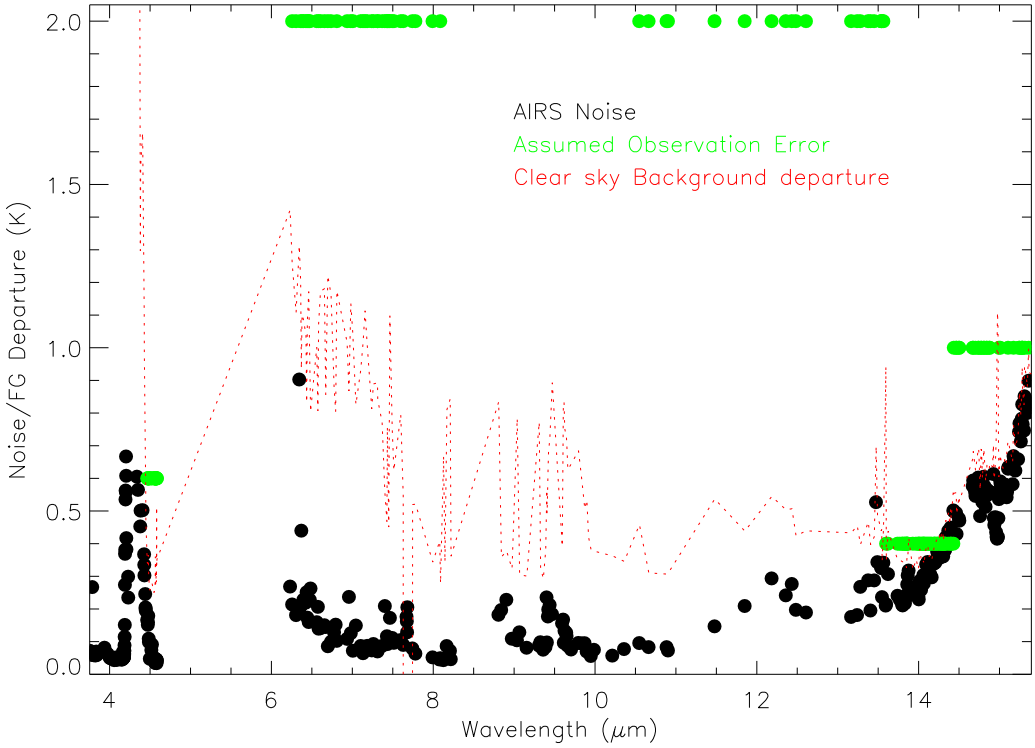


Figure 7: A comparison of the AIRS instrument noise and the assumed observational noise in the assimilation of AIRS radiances.

NOAA/NESDIS supply a reconstructed radiances data set — also in near-real-time — for the same 324 channels as the real radiances. The radiances are calculated using the first 200 principal components of the climatological covariance of *observed* AIRS radiances (taken from the full set of observed data, i.e., clear and cloudy scenes). The principal components are derived from 1688 of the 2378 channels, as a number of noisy "popping" channels are omitted together with those channels around 4.3μm which are highly affected by non-LTE effects during the daytime. All data comes with a quality control flag which indicates whether the original spectrum has been reconstructed to within the noise limits.

The effect of the noise smoothing is most clearly illustrated in the stratospheric channels in the 15μm CO_2 band where the instrument noise is much larger than the other noise terms (at least for the real radiances) and the brightness temperature fields are intrinsically smooth. Figure 8 compares real and reconstructed radiances for a single granule (six minutes) of AIRS data for AIRS channel 101 at 674.41cm^{-1}. The apparent spatial smoothing is purely a result of spectral smoothing within each field-of-view.

The importance of the quality controlling the reconstruction process is illustrated in Figures 9 and 10 with an extreme example from 1st March 2005. Figure 9 shows the difference between the reconstructed and real radiances for AIRS channel 72 (667cm^{-1}). Differences are within $\pm$1K except for one scan line where differences are up to 14K. Figure 10 compares the spectra for one of these anomalous fields of view and also spectra for a field of view from an adjacent scan line. The reconstructed spectrum for the anomalous field of view is clearly very different to the original although a similar spectrum in the adjacent scan line is reconstructed well. The reason for this anomaly could be that the original spectrum fell outside of the range of situations used to derive the principal components (either because of an unusual atmospheric state or a calibration problem) or some other processing issue. Most important is that the discrepancy was flagged through the monitoring of the reconstruction scores (by the data provider) and so the observation can be rejected.

An initial assimilation experiment using the reconstructed radiances but using the same assumed observation noise as is used with real AIRS radiances was conducted for the period from 1st March 2005 – 30th April 2005.

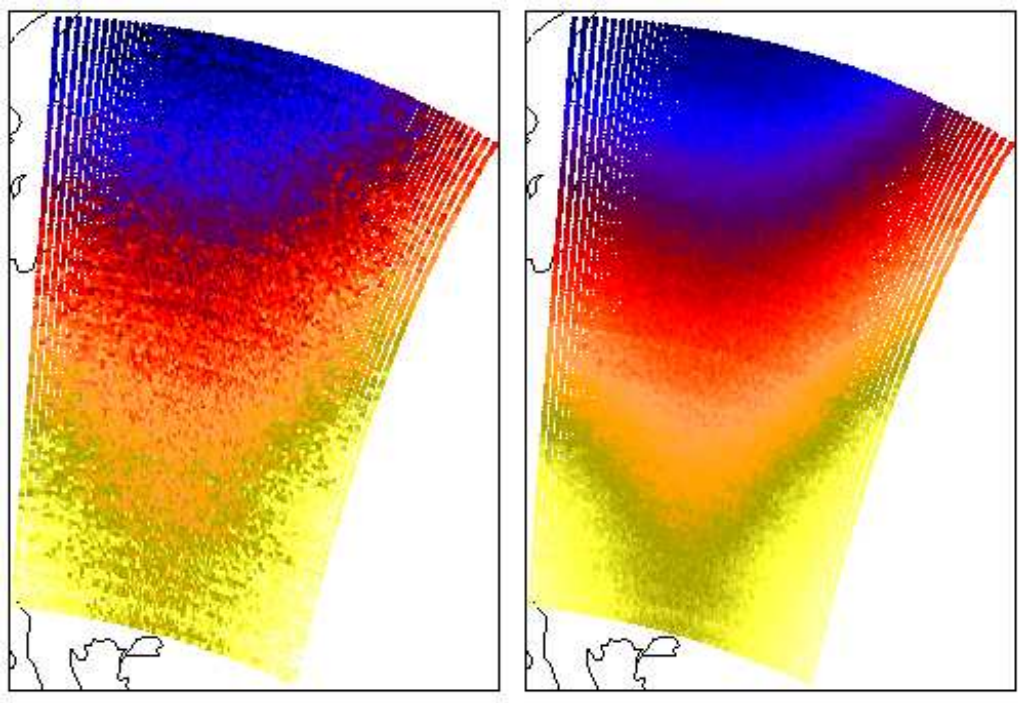


Figure 8: A single granule (six minutes) of AIRS data before (left) and after (right) noise smoothing through the reconstructed radiances approach. The plots are for AIRS channel 101 at 674.41cm^{-1}.

Figure 11 shows a comparison between the standard deviation of the background departures (the differences between the observed radiances and radiances calculated from the short-range (3–15hr) forecast fields — the background — used as *a priori* information in the assimilation process) for real and reconstructed radiances for the first cycle of this experiment (so that the background model field is the same for the two cases). The standard deviations of the reconstructed radiances' background departures are significantly reduced for all those channels where instrument noise is expected to be a significant contributor to the total departure. The biases (Figure 12) are unchanged to within 0.1K, as the effect of reconstruction is to smooth random noise rather than change the mean radiances.

As explored in the previous section, one expects the errors in the reconstructed radiances to be correlated between channels. To verify this we can compare the covariance matrix of the background departures of used data for real radiances with that for reconstructed radiances. The result for the first 120 AIRS channels in the 324 channel subset is shown in Figures 13 and 14. For real radiances the departures are generally uncorrelated except where vertically correlated background errors in the stratosphere (e.g., around channel 35) are significant. It can be seen that the reconstruction process introduces significant additional correlations into the departures.

The effect of using reconstructed radiances can be seen on the process of cloud detection. The ECMWF cloud detection scheme (McNally and Watts, 2003) ranks the AIRS channels according to the lowest height in the atmosphere to which they are sensitive. It then looks for a consistent (usually negative) signal in the first guess departures that would indicate where in the vertical a cloud is likely to be. As cloud (and model) errors are expected to vary slowly between channels whereas instrument noise is expected to vary rapidly, a filter with a window width of typically ten channels is applied to reduce the effect of instrument noise on the algorithm. In Figures 15, the cloud detection algorithm is illustrated for an observation (over the Antarctic plateau) for real and reconstructed radiances. The higher noise in the real radiances is apparent (although it should be remembered that similar channels will have correlated noise in the reconstructed case, further reducing the high-frequency noise in this figure). The filtered signal is affected so that five channels are flagged as cloudy for the reconstructed radiances that are flagged clear for the real radiances. Despite the above illustration, the

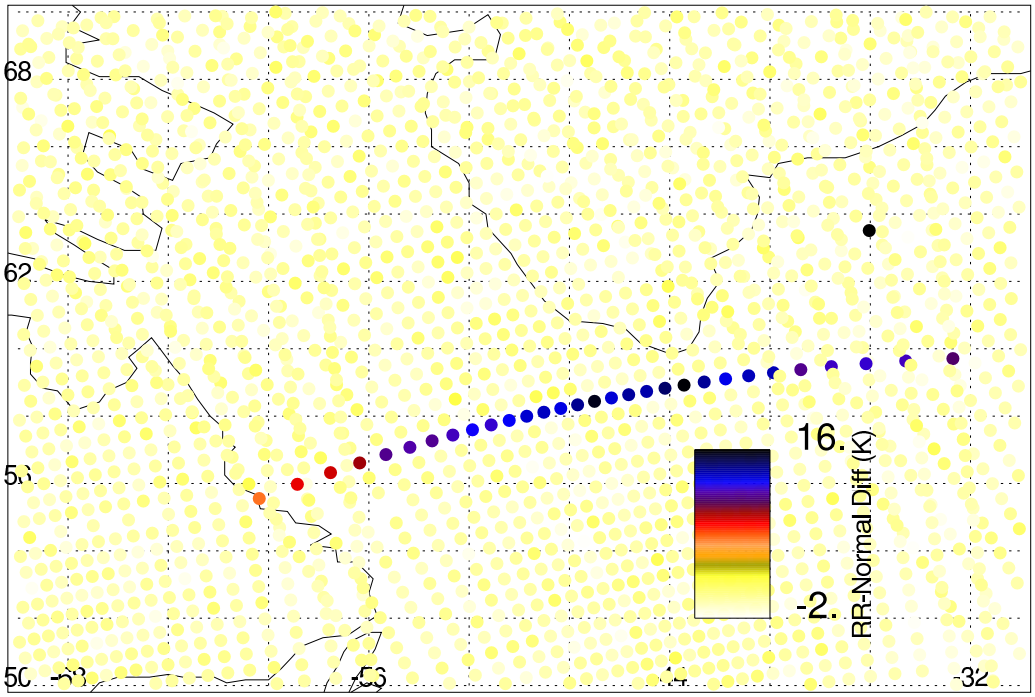


Figure 9: A map of the difference between the real and reconstructed observed brightness temperatures for AIRS channel 72 illustrating a case where the reconstruction process has introduced large errors. These cases are identified and flagged through the monitoring of reconstruction scores.

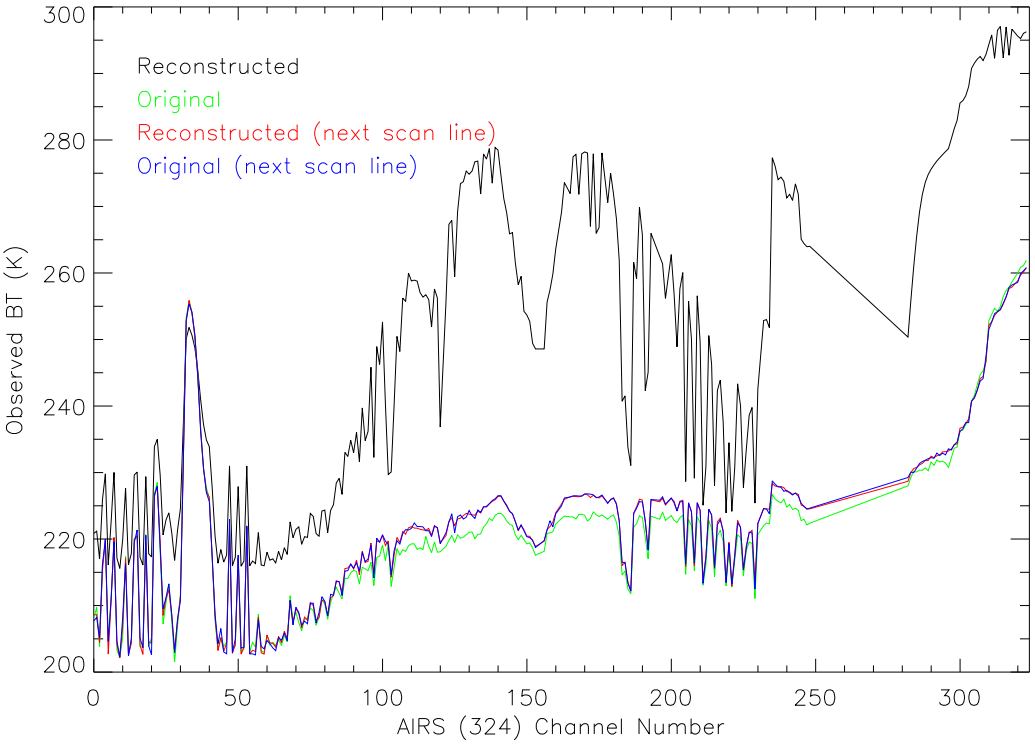


Figure 10: Reconstructed and real spectra for adjacent fields of view where the reconstruction process has worked well and where it has introduced errors.

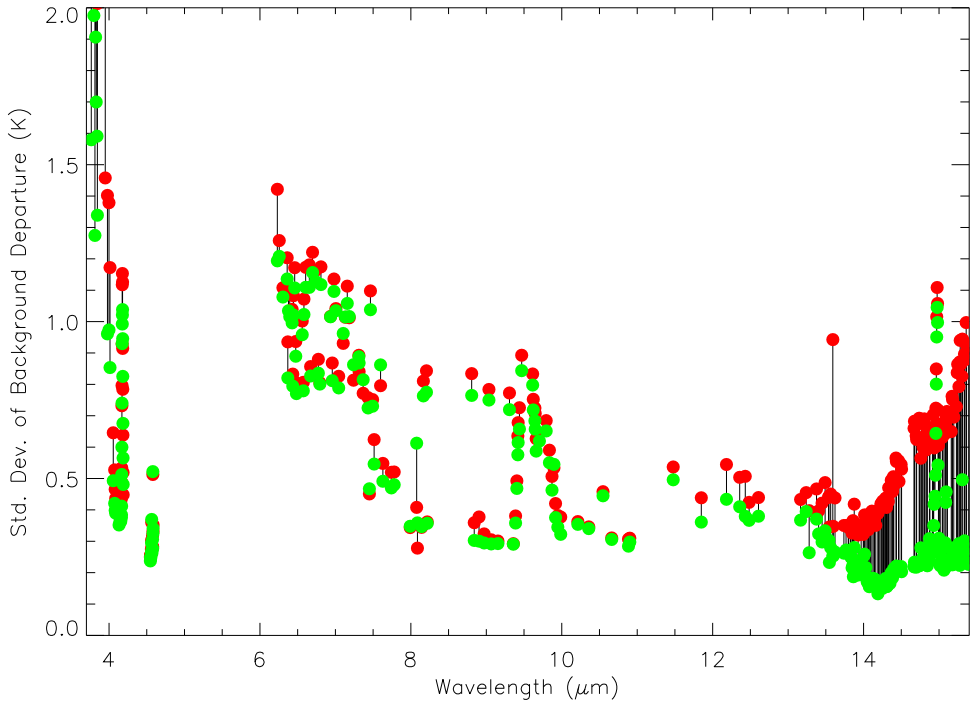


Figure 11: A comparison of the standard-deviations of clear-sky departures from the same model background for real (red) and reconstructed (green) radiances. Significant "denoising" is seen in the 15μm band where instrument noise is dominant over model error. The apparent increase in departure standard deviation for the channel at 8.07μm is an artifact arising from the cloud detection scheme.

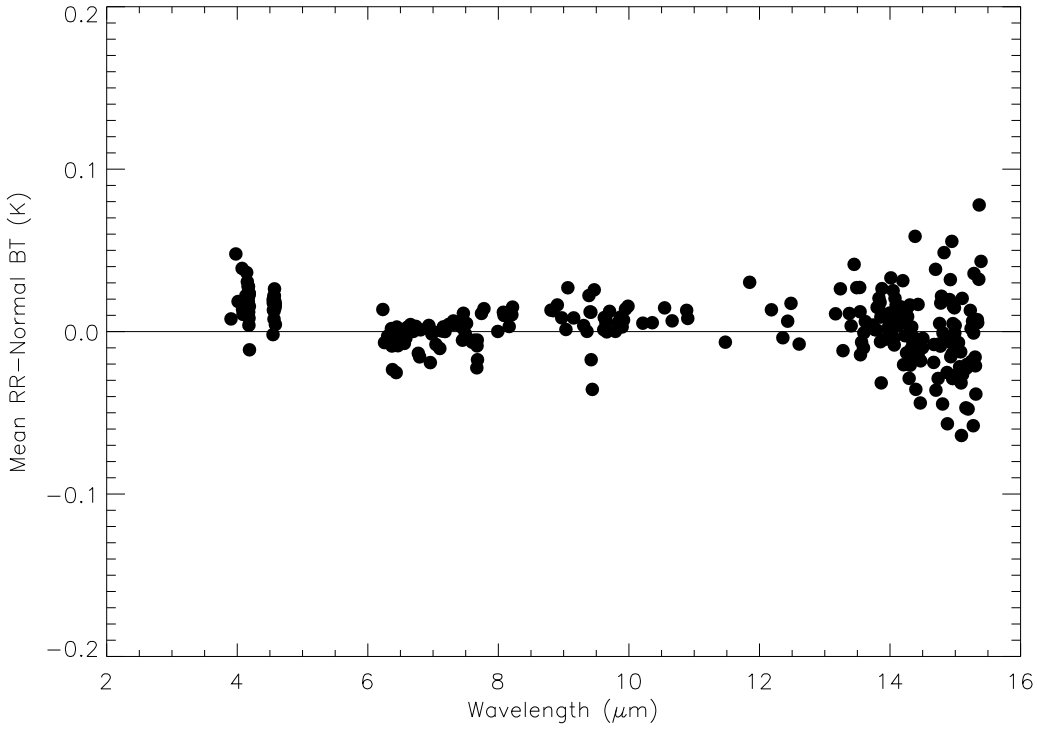


Figure 12: The difference between the mean clear-sky departures from the same model background for real and reconstructed radiances.

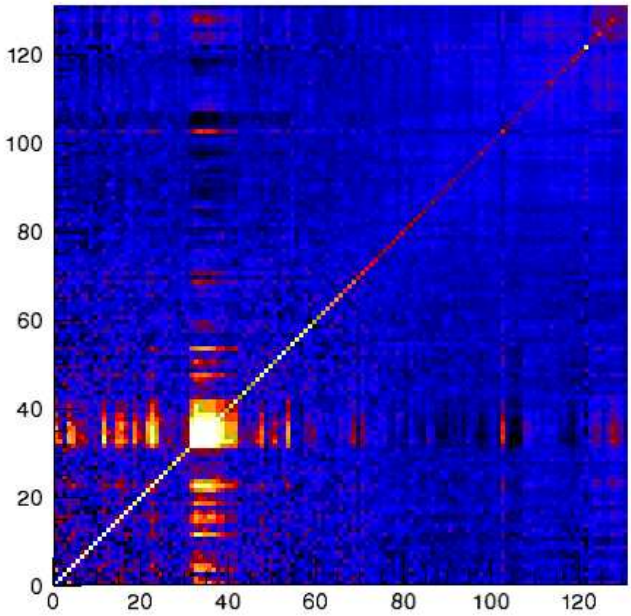


Figure 13: The covariance matrix of the background departures for the first 120 channels of the 324 channel subset (covering the region 649–739cm^{-1}) for real radiances. The instrument noise is seen to be very close to diagonal while off-diagonal covariances are only significant for channels that sound the stratosphere - where the background error outweighs the instrument noise.

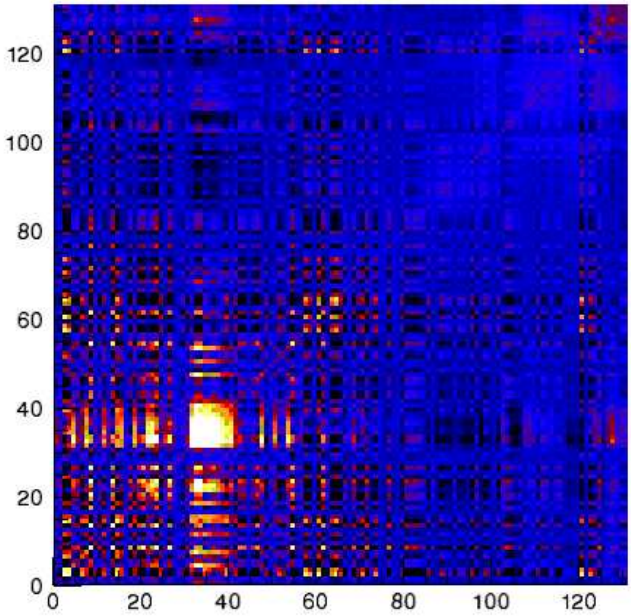


Figure 14: As Figure 13 except for reconstructed radiances, the colour scale is the same. The diagonal noise is seen to be greatly reduced where instrument noise is dominant but this noise is now correlated between similar channels.

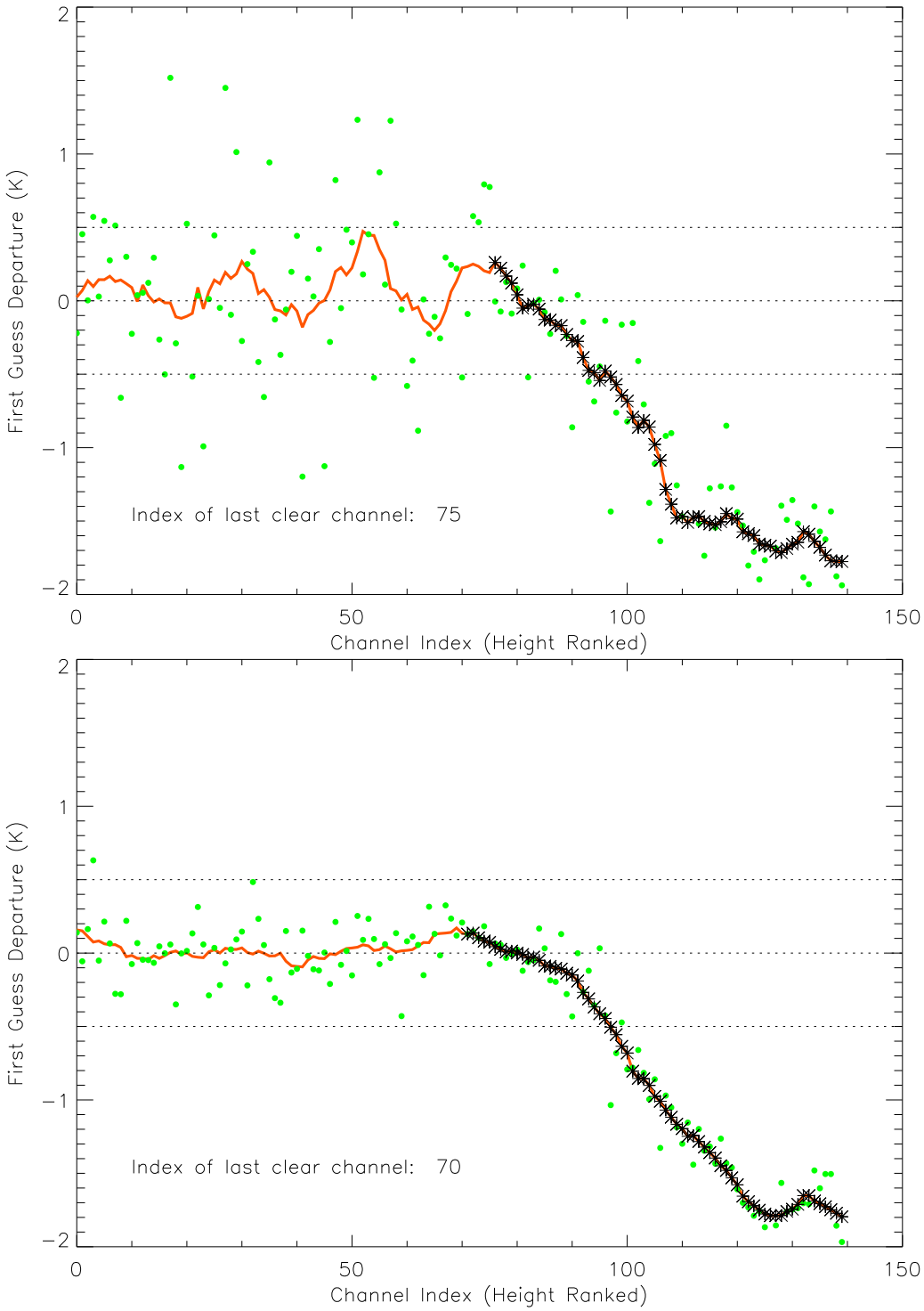


Figure 15: The cloud detection algorithm illustrated for the same observation but for real (top) and reconstructed (bottom) radiances The green dots are the raw first guess departures, the orange line is the filtered radiances and the asterisks show the channels that have been flagged as cloudy. The noise reduction for the reconstructed radiances is clear and five channels are flagged as cloudy in the reconstructed radiances which are clear for the real radiances. These channels are all from the 10.2–15.4µm band.

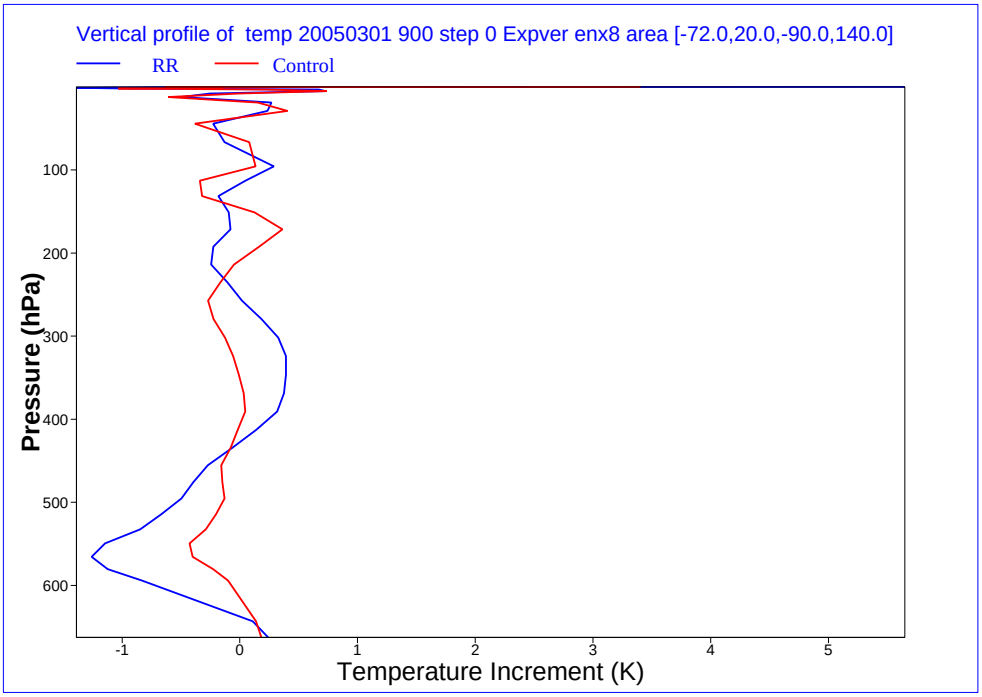


Figure 16: Mean temperature increments above the Antarctic plateau during March and April 2005 of using real and reconstructed radiances. The use of reconstructed radiances damps the oscillations seen in the stratosphere but increases the anomalous near surface increment caused by contamination from the surface.

robustness of the cloud detection algorithm is such that typically less than 2% of channels are flagged as cloudy for reconstructed radiances while clear for real radiances and less than 1% differ in the opposite sense.

Temperature increments are largely unchanged on switching from real to reconstructed radiances apart from in the polar regions. Figure 16 shows erroneous oscillatory increments (known to have been a feature of the ECMWF analysis and since largely removed through the assimilation of GPSRO data and improvement of the radiative transfer for AMSU-A channel 14) in the polar stratosphere are damped using reconstructed radiances. It appears that the reconstructed radiance process has denoised the radiances and effectively reduced the "null space" in which the oscillations occur. However, the reconstructed radiances have also reinforced an increment near the surface believed to be related to surface contamination of the AIRS observations. The low-level degradation appears to be related to ten channels being used above the high antarctic plateau that have a significant contribution from the surface (which is not accurately modelled). Removal of these low-level channels over Antarctica results in a mean profile over the Antarctic plateau in much closer agreement to that when real radiances are used (Figure 17).

The effect of this change in analysis increments can also be seen in the agreement of the assimilation with Antarctic radiosondes (18) which is improved on using reconstructed rather than real radiances for levels above 250hPa, but is somewhat degraded at lower levels. Removal of the low level channels over Antarctica results in an improvement in the agreement of radiosondes and model in the region around 500hPa (19).

Apart from the above and the noise reduction of the AIRS channels that has already been presented, the fit to other observations is largely unaffected by the move from real to reconstructed radiances.

The impact on the Southern Hemisphere 500hPa geopotential height forecast anomaly correlation scores is shown in Figures 20 (before the removal of the ten channels above) and 21 (after these channels have been removed). The effect on the Southern Hemisphere forecast scores on switching to reconstructed radiances is neutral (although if the surface contaminated channels are not removed there is a small negative impact).

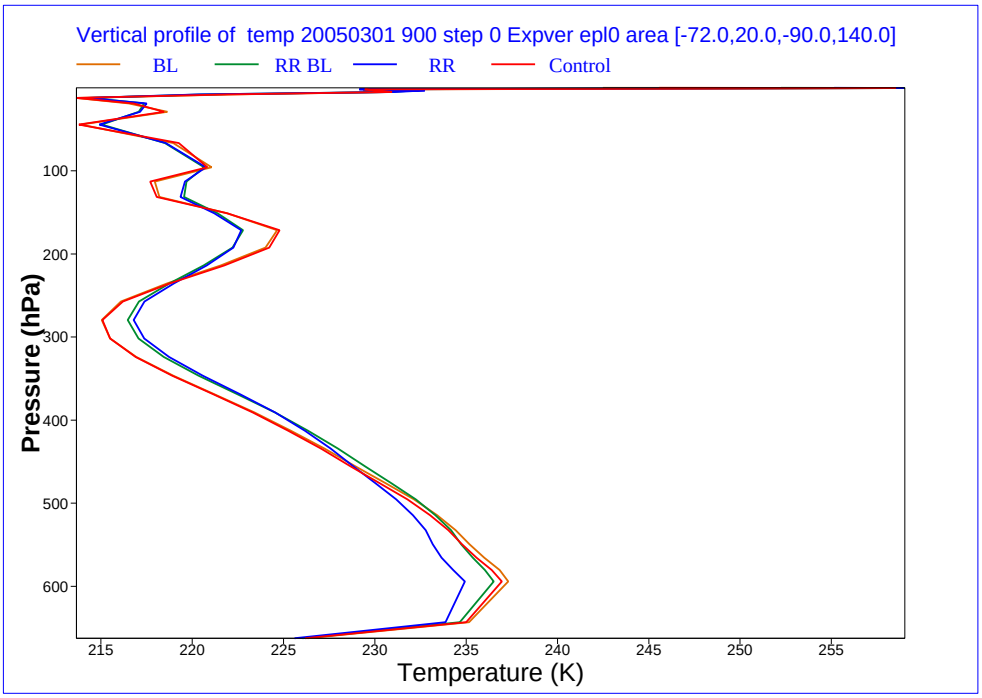


Figure 17: The effect on the mean temperature profile above the Antarctic plateau during March and April 2005 of using real and reconstructed radiances and also of blacklisting ten channels that were being affected by the surface. The main effect of the reconstructed radiances is to damp the spurious stratospheric oscillation and to introduce an erroneous cooling in the near surface layers. The additional blacklisting reverses this cooling.

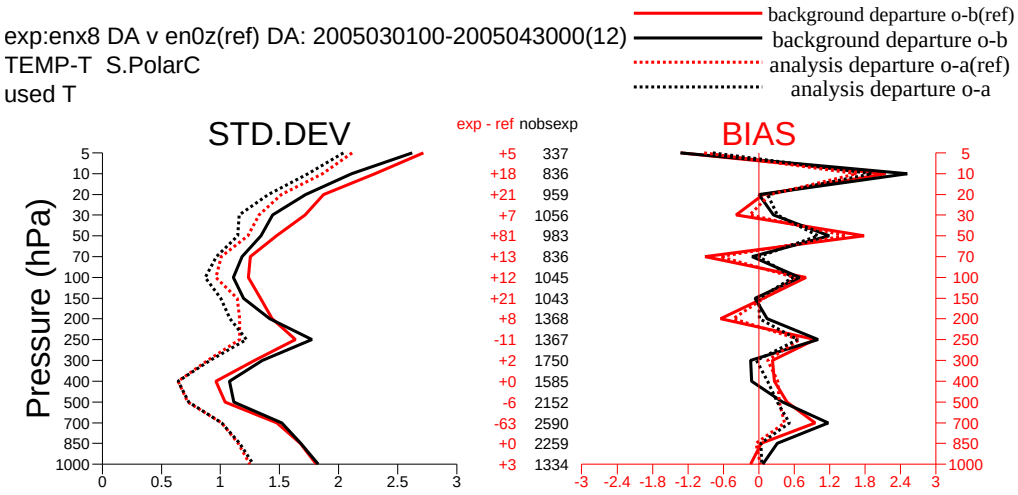


Figure 18: Mean and standard deviations of analysis and background departures for radiosondes around Antarctica. The reference curves (red solid and cyan dotted) are for real radiances and while the experiment curves (black solid and green dotted) are for reconstructed radiances. The use of reconstructed radiances appears to remove some of the spurious oscillations in the departures for levels above 250hPa.

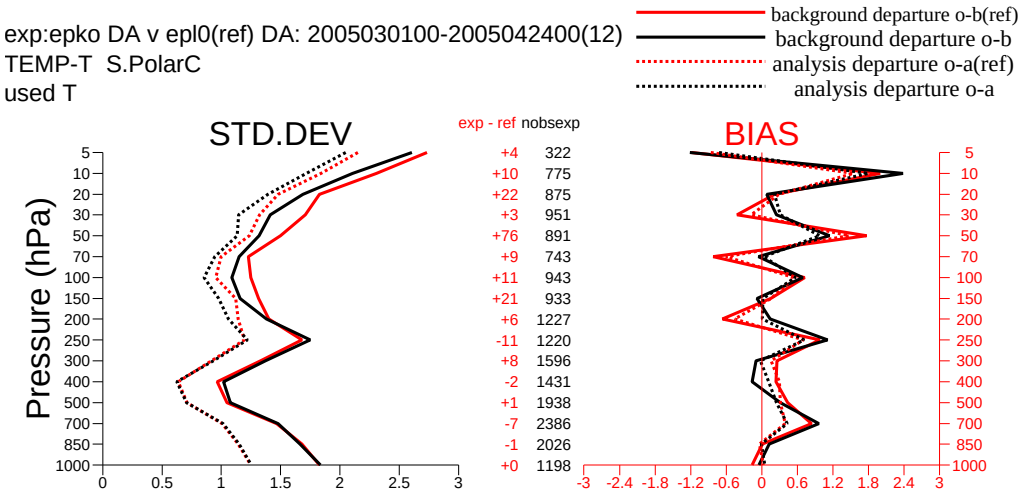


Figure 19: As Figure 18 but with ten AIRS channels that are affected by the surface of the high antarctic plateau removed. Note the improvement in the lower level departures around 500hPa.

The impact is also neutral in the Northern Hemisphere. It is important to reiterate that these impacts are relative to the real radiance assimilation and the impact is still very positive in the Southern Hemisphere using reconstructed radiances compared to using no AIRS data.

The results of a simple calculation illustrating the issues surrounding the choice of observation noise is shown in Figure 22. This is an example linear analysis of retrieval errors illustrating the effect of assuming the wrong observational error covariance matrix (following the method of Watts and McNally, 1988) for the 155 AIRS channels used at ECMWF. Watts and McNally show that if the theoretical analysis error covariance calculated using the assumed observation error covariance matrix, $\mathbf{E}$, is $\mathbf{A}$ then if the true observation error covariance is $\mathbf{E}'$ the actual analysis error covariance, $\mathbf{A}'$ is given by:

$$\mathbf{A}' = \mathbf{A} + \mathbf{W}(\mathbf{E}' - \mathbf{E})\mathbf{W}^T \tag{11}$$

where $\mathbf{W} = \mathbf{B}\mathbf{H}^T(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{E})^{-1}$, $\mathbf{B}$ is the background error covariance matrix and $\mathbf{H}$ is the Jacobian matrix.

In the figure, the background error is shown by the red line, while the retrieval using the real (unreconstructed) radiances is shown in black. The cyan curve shows the expected retrieval error if the assumed errors are kept the same but reconstructed radiances are used — not a large difference — while the green curve shows that a much larger impact is obtained when the correct, correlated errors are assumed for the reconstructed radiances. Finally, if just the diagonals of the reconstructed radiances' covariance matrix are assumed, the result can even be a degradation (blue curve). Therefore, for the maximum impact of the reconstructed radiances to be obtained one must fully account for the correlations in the error covariance matrix.

Experiments reducing the assumed observation errors have yielded neutral or negative results in terms of fit to other observations and forecast scores. Similarly initial encouraging results on using the calculated reconstructed radiances' correlation structure on top of the current operational assumed error covariance matrix have yielded essentially neutral results.

Attempts have been made to do a comparison between real and reconstructed radiances using full error covariance matrices containing best estimates of the various error terms involved. The terms considered were instrument noise, forward model noise, undetected cloud and "non-linearity" error (the error introduced on assimilation through the true and analysis Jacobians being different). The instrument noise and its transformation from real to reconstructed radiances can be estimated very well, but the remaining terms (especially the random component of the forward model error) are hard to determine. Furthermore, there are the additional compli-

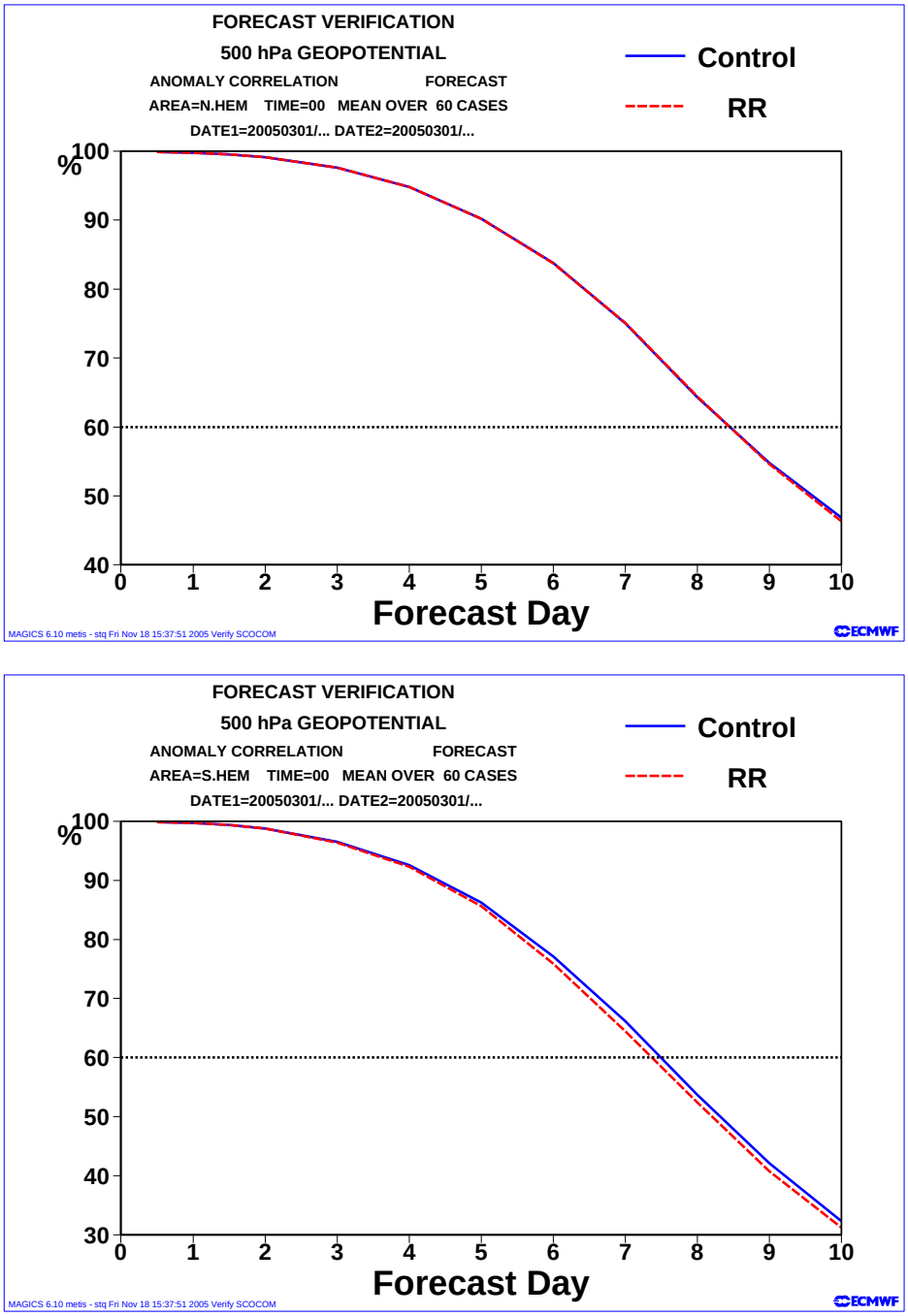


Figure 20: Anomaly correlation scores for Southern Hemisphere 500hPa geopotential height forecasts for 1st March 2005 – 30th April 2005. There is a small degradation when replacing real radiances with reconstructed radiances. The Northern Hemisphere is neutral.

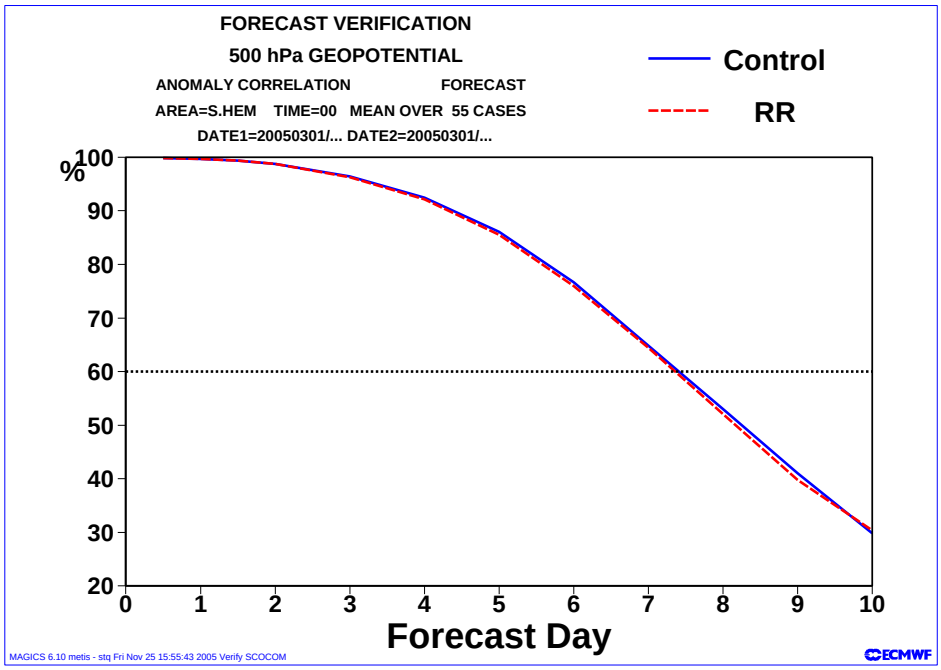


Figure 21: Southern Hemisphere anomaly correlation scores for 500hPa geopotential height forecasts for 1st March 2005 – 24th April 2005 where ten AIRS channels that are affected by the surface of the high Antarctic plateau are removed where the terrain is higher than 2000m. The negative impact of the reconstructed radiances relative to real radiances seen in the initial run is greatly reduced.

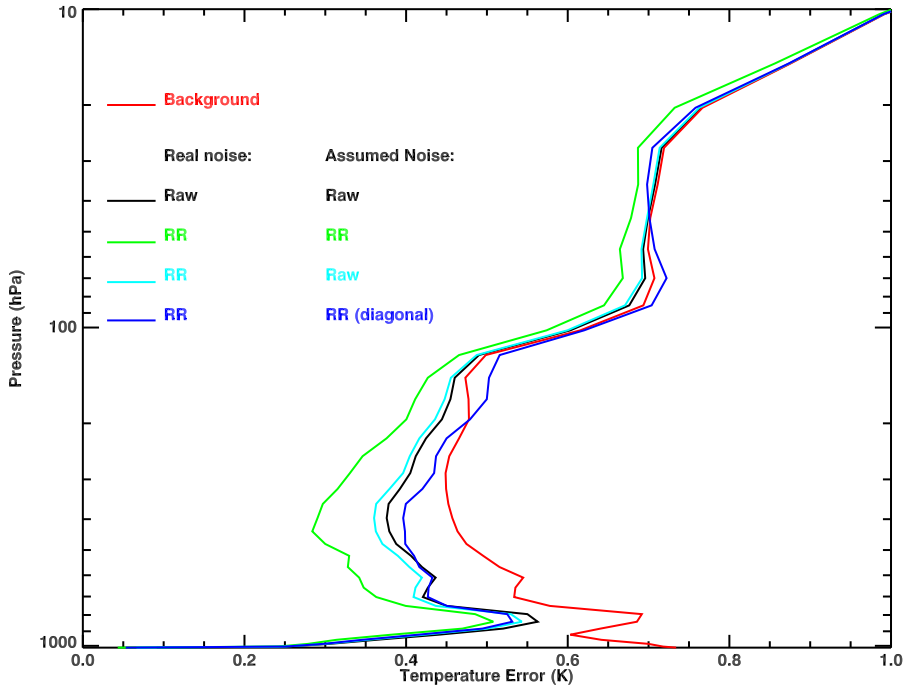


Figure 22: A simple linear analysis of the effect of using the wrong observational error covariance matrix with reconstructed radiances.

cations of spatially correlated error (especially from forward model error) and representivity error which were not addressed. These attempts did not improve on the current operational system and are not presented here but it is recognised that the proper characterisation of all these error sources is an important continuing process in the optimal assimilation of these radiances — whether reconstructed or not.

Another possibility, as yet unexplored, for the use of radiances with correlated errors is that channels are selected such that they have minimal mutual error correlation, thus allowing a more realistic diagonal error covariance matrix to be used.

In conclusion, the AIRS experience has shown that the reconstructed radiances are a robust product, with small differences in the mean but marked reduction in the instrument noise. The introduction of inter-channel correlations in the instrument noise has been observed in departure statistics. Significant impact on the assimilation has been limited to polar regions where some improvements were seen but also a degradation where the data were being used inappropriately. The impact on forecast scores is essentially neutral (relative to real radiances — the impact is very positive relative to the complete removal of AIRS from the system).

5 Discussion and Conclusions

Principal component analysis and reconstructed radiances are two techniques that can be used to compress the full spectrum from a kilochannel infrared sounding instrument into a few hundred "channels".

The direct assimilation of principal component amplitudes requires, due to the non-local Jacobians associated with the principal components, either that only clear columns are assimilated or that cloud properties be included in the analysis — something that is still only at the research phase in current NWP systems. The non-locality of Jacobians is also a practical issue in that, for example, they can result in unwanted influence of the stratosphere on the tropospheric analysis — something that can conceivably be corrected by using appropriate channels when generating the principal components.

The use of reconstructed radiances is potentially a more promising approach to representing the spectra in compressed form as the reconstructed radiance channels have very similar Jacobians to normal radiances. Initial experiments with reconstructed radiances for AIRS show that the properties of the radiances are different to normal radiances (in terms of the radiances themselves and their influence on model fields) but that the impact on the forecast scores is very similar.

The questions that remain with these datatypes is whether it is possible to extract further forecast impact above and beyond what has already been demonstrated. It should first be noted that it has not been demonstrated that the effect of adding the information in significantly more channels than are already assimilated will have significant positive impact on model fields. Indeed, it can be said for IASI (and to a lesser extent for AIRS) that the number of channels assimilated operationally into NWP models is not far below the maximum for the important 15μm CO_2 band. The bulk of the remainder of the spectrum is made up of water vapour and surface sensitive channels that are challenging to assimilate in bulk for reasons beyond the scope of this discussion.

If, however, the additional information in the remainder of the spectrum can improve forecast skills, how can one use this information in the context of principal components or reconstructed radiances? The accurate specification of the observation error covariance matrix (including terms from the forward model and due to representivity) will be very important.

Related to this is having a background state that is as unbiased as possible plus an accurate bias correction to allow for forward model and instrument biases. The importance of the bias is increased as the random component of the noise is decreased due to the noise-smoothing properties of the PCA method. Finally, as the assumed instrument noise is decreased (and therefore the weight given to the observations in the assimilation system is increased) the ability of erroneous observations to negatively affect the analysis is increased. Quality control of the observations (including cloud detection) therefore becomes ever more important.

In summary, the principal component approach allows the information contained in the full spectrum of a kilochannel infrared sounder to be represented in a few hundred numbers. This additional information is potentially useful in improving our knowledge of the atmospheric state and thus improving forecast skill.

Acknowledgements

I would like to thank the following for help and useful discussions in the preparation of this paper: Sean Healy, Fiona Hilton, Tony McNally, Marco Matricardi, Nigel Atkinson and Tim Hultberg. Parts of this work were under EUMETSAT Contract No. EUM WP 989-2 and through the EUMETSAT NWPSAF.

References

Aires, F., W.B. Rossow, N.A. Scott and A. Chédin (2002a). Remote sensing from the infrared atmospheric sounding interferometer instrument. 1. Compression, denoising, and first-guess retrieval algorithms. *J. Geophys. Res.*, **107(D22)**, 10.1029/2001JD000955.

Aires, F., W.B. Rossow, N.A. Scott and A. Chédin (2002b). Remote sensing from the infrared atmospheric sounding interferometer instrument. 2. Simultaneous retrieval of temperature, water vapor, and ozone atmospheric profiles. *J. Geophys. Res.*, **107(D22)**, 10.1029/2001JD0001591.

Antonelli, P., H.E. Revercomb, L.A. Sromovsky, W.L. Smith, R.O. Knuteson, D.C. Tobin, R.K. Garcia, H.B. Howell, H.-L. Huang, and F.A. Best (2004). A principal component noise filter for high spectral resolution infrared measurements, *J. Geophys. Res.*, **109**, D23102–23124.

Auligné, T., F. Rabier, L. Lavanant, and M. Dahoui (2003). First results of the assimilation of AIRS data in Météo-France Numerical Weather Prediction model. *Proc. of the 13th International TOVS Study Conference, St. Adele, Canada, October 2003*

Chalon, G., F. Cayla, and D. Diebel (2001). IASI: An Advanced Sounder for Operational Meteorology. *Proc. 52nd Congress of IAF, Toulouse France, 1-5 Oct. 2001*. http://smsc.cnes.fr/IASI/A_publications.htm

Collard, A.D., R. Saunders, J. Cameron, B. Harris, Y. Takeuchi, L. Horrocks (2003). Assimilation of data from AIRS for improved numerical weather prediction. *Proc. of the 13th International TOVS Study Conference, St. Adele, Canada, October 2003*

Collard, A.D., and A.P. McNally (2009). The assimilation of Infrared Atmospheric Sounding Interferometer radiances at ECMWF. *Q.J.R. Meteorol. Soc.*, **135**, 1044–1058.

S. Havemann, J.-C. Thelen, J.P. Taylor, A. Keil (2009) The Havemann-Taylor Fast Radiative Transfer Code: Exact fast radiative transfer for scattering atmospheres using Principal Components (PCs). *Current Problems in Atmospheric Radiation (IRS 2008): Proceedings of the International Radiation Symposium (IRC/IAMAS). AIP Conference Proceedings*, **1100**, 38–40.

Hilton, F., N.C. Atkinson, S.J. English and J.R. Eyre (2009a). Assimilation of IASI at the Met Office and assessment of its impact through observing system experiments. *Q.J.R. Meteorol. Soc.*, **135**, 495–505.

Hilton, F., A. Collard, V. Guidard, R. Randriamampianina, and M. Schwaerz (2009b). Assimilation of IASI Radiances at European NWP Centres. *Proc. of the ECMWF/EUMETSAT NWP-SAF Workshop on the assimilation of IASI in NWP, ECMWF, Reading, UK, 6th – 8th May 2009*

Huang, H.-L., W.L. Smith and H.M. Woolf (1992). Vertical resolution and accuracy of atmospheric infrared sounding spectrometers. *J. Appl. Meteor.*, **31**, 265–274.

Huang, H.-L. and P. Antonelli (2001). Application of principal component analysis to high-resolution infrared

measurement compression and retrieval. *J. Appl. Meteor.*, **40**, 365–388.

Kelly, G. and J.-N. Thépaut (2007). Evaluation of the impact of the space component of the Global Observing System through Observing System Experiments, *ECMWF Newsletter. No. 112 Autumn 2007.*

Lee, A.C.L., and S. Bedford (2004). Support study on IASI Level 1c Data Compression. *Report EUMETSAT Contract EUM/CO/03/1155/PS, Met Office, Exeter. 149pp.*

Le Marshall, J., J. Jung, J. Derber, M. Chahine, R. Treadon, S.J. Lord, M. Goldberg, W. Wolf, H.C. Liu, J. Joiner, J. Woollen, R. Todling, P. van Delst, and Y. Tahara (2006). Improving Global Analysis and Forecasting with AIRS. *Bull. Amer. Meteor. Soc.*, **87**, 891-894.

Liu, X., W.L. Smith, D.K. Zhou and A.M. Larar (2005). A principal component-based radiative transfer forward model (PCRTM) for hyperspectral instruments. *Proceedings of SPIE — Vol. 5655. Multispectral and hyperspectral remote sensing instruments and applications II, Allen M. Larar, Makoto Suzuki, Qingxi Tong, Editors,* pp. 96–105.

Matricardi, M. (2009). An observation operator for the assimilation of principal component scores into a NWP system. *Report EUMETSAT Contract EUM/CO/07/4600000475/PS, ECMWF, Reading, UK. 64pp.*

McNally, A.P., P.D. Watts, J.A. Smith, R. Engelen, G.A. Kelly, J.-N. Thépaut and M. Matricardi (2006). The assimilation of AIRS radiance data at ECMWF. *Q.J.R. Meteorol. Soc.*, **132**, 935–957.

McNally, A.P. and P.D. Watts (2003). A cloud detection algorithm for high-spectral-resolution infrared sounders. *Q.J.R. Meteorol. Soc.*, **129**, 2411–2323.

Smith, W.L., D.K. Zhou, X. Liu, H.-L. Huang, H.E. Revercomb, A.M. Larar, and C.D. Barnet (2005). Ultra high spectral resolution satellite remote sounding — Results from aircraft and satellite measurements. *Proc. of the 14th International TOVS Study Conference, Beijing, May 2005*

Tobin, D.C., P. Antonelli, H.E. Revercomb, S. Dutcher, D.D. Turner, J.K. Taylor, R.O. Knuteson, and K. Vinson (2007) Hyperspectral data noise characterization using principle *[sic]* component analysis: application to the atmospheric infrared sounder. *J. Appl. Remote Sens.*, **1**, 013515.

Turner, D.D., R.O. Knuteson, H.E. Revercomb, C. Lo and R.G. Dedecker (2006). Noise reduction of Atmospheric Emitted Radiance Interferometer (AERI) observations using principal component analysis". *J. Atmos. Oceanic Technol.*, **23**, 1123–1238.

Watts, P.D., and A.P. McNally (1988). The sensitivity of a minimum variance retrieval scheme to the values of its principal parameters. *Technical Proceedings of the Fourth International TOVS Study Conference.*, 399–407.