The Assimilation of AIRS radiance data at ECMWF

A.P. McNally, P.D. Watts, J.A. Smith, R. Engelen, G. Kelly,
J.N. Thépaut, and M. Matricardi

ECMWF
Shinfield Park, Reading, UK

Summary

The development of an assimilation system for radiance data from the Atmospheric InfraRed Sounder (AIRS) is described, in particular, the identification of cloud contamination, bias correction and the characterization of errors in the measured radiances and radiative transfer model. The results of assimilation experiments are presented. These show that a conservative use of AIRS radiance data (in a system already extensively observed with other satellite data) results in a small, but consistent improvement in the quality of analyses and forecasts. Larger impacts of AIRS are found in hypothetical experiments that test the use of radiances from only a single sounding instrument. In these, the AIRS is found to outperform the use of data from either a single Advanced Microwave Sounding Unit (AMSUA) or from a single High-resolution InfraRed Sounder (HIRS). In this hypothetical context the relative forecast performance of each sensor is found to correlate with the size and vertical scale of increments caused by the assimilation of the radiances.

1. Introduction

The NASA AQUA spacecraft was launched in May 2002 and a subset of radiance data from the AIRS instrument (a grating spectrometer with 2378 channels covering the thermal infrared spectrum between 3 and 15 microns, Auman et al. 2003.) has been available in near-real-time (NRT) since the end of October 2002. The sampled data consists of 324 channels, manually selected to represent the main features of the observed spectra. Spatially, the data are thinned to provide radiances from only one AIRS spot in every nine (i.e. the central spot from a 3x3 array).

While the AIRS instrument is formally part of a NASA research mission, it has received a significant amount of attention from operational Numerical Weather Prediction (NWP) agencies. The primary reason for this is that it provides an advanced look at the type of data we expect from the next generation of infrared sounders soon to become key elements on operational satellite missions (e.g. IASI on the European METOP-1 and CrIS on the American NPOESS).

High spectral resolution sounders such as the AIRS are capable of providing information on atmospheric temperature and composition at a higher vertical resolution than is achievable with the filter infrared radiometers carried by current operational satellites. Indeed simulations (such as those reported in Prunet et al. 1998) have suggested that the assimilation of these data could help constrain small-scale baroclinic structures known to be important in the development of forecast errors. However, while simulations provide a useful measure of information content, they are inevitably optimistic and take no account of practical difficulties that detract from the utility of real data in a real assimilation system (for example the presence of cloud, McNally 2002, Fourrie and Rabier 2004).

2. Evaluation of the AIRS radiance data and RTM

Before any assimilation experiments were performed the degree of consistency between AIRS radiances simulated by the RTM and the radiances measured by the instrument was evaluated. This is essential, as it not only quantifies the accuracy of the instrument calibration and spectral characterization, but also the accuracy of the RTM. The RTM used is an adaptation of the RTTOV model (version 6 as described in Matricardi et al. 2001). It is a fast model in that it uses regression to estimate transmittances from atmospheric parameters, the regression being previously trained on a set of accurate line-by-line calculations. In this study the RTM is supplied with a fixed surface emissivity of 0.98 over all surfaces. Figure 1 shows global mean observed minus background radiance departures for data determined as clear (the identification of clear data will be discussed in the next section) with, for clarity, the spectrum split in to 4 bands. The statistics have been evaluated over a one-month period (August 2003) to ensure that channels that are frequently removed because of cloud contamination are adequately represented. For channels sensitive to the surface only sea points are used in the statistics. In the long-wave temperature sounding band (figure 1a) biases are generally small (typically less than 1K) with a distinct transition from positive to negative values in the more transparent channels (that latter possibly reflecting some residual cloud contamination). Larger bias values are found in channels around 15 microns. These are due to known systematic errors in the background estimates of upper stratospheric / mesospheric temperatures (to which these channels are very sensitive). The ozone sounding channels around 9.5 microns (figure 1b) show a curious bias structure thought to be associated with spectroscopy errors. The AIRS channels sensitive to low level water vapour (left side of figure 1c) are positively biased, whereas the strong water vapour (right side of figure 1c) are negatively biased. The latter may be due to errors in the RTM, but is more likely a result of the background upper tropospheric humidity being systematically too dry (as a very similar signal is seen in the departure statistics of HIRS and AMSUB radiance data). The largest biases are seen in the AIRS shortwave channels of figure 1d. The statistics (computed only for night) show the tropospheric channels around 4.5 microns having acceptable values (comparable in size to the long-wave), but very large biases (as yet unexplained) are found in the stratospheric channels near the band centre (4.3 microns) and in the very transparent

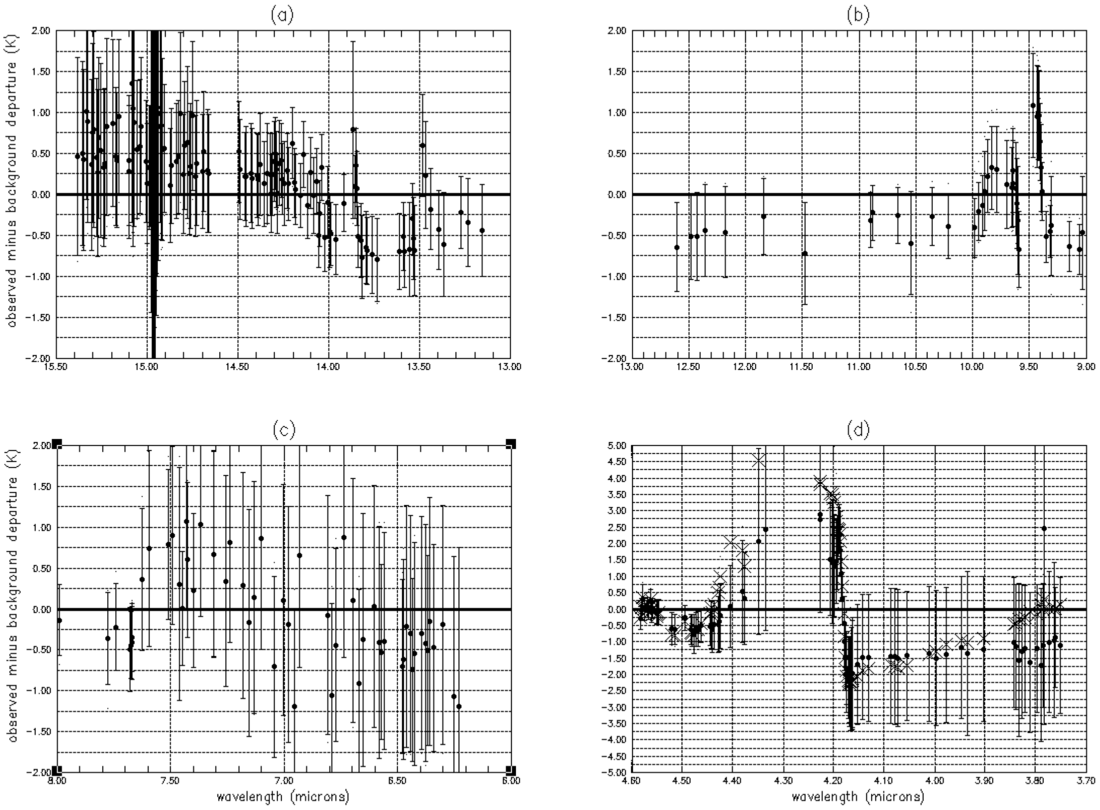

Figure 1 Globally averaged observed minus background departures for (a) the AIRS long-wave sounding band, (b) window channels and ozone channels, (c) the AIRS water vapour band and (d) the short-wave band (night data only). The black dots show the mean value and the length of the vertical bar is equal to twice the standard deviation. All statistics were evaluated during August 2003 for clear data and the units are degrees Kelvin.

channels towards shorter wavelengths. The daytime statistics for the shortwave band (not shown) are even worse. Stratospheric shortwave observations demonstrate effects consistent with the atmosphere not being in local thermodynamic equilibrium (LTE) (an effect not modelled in the RTM) and are consequently significantly biased relative to computed values. Tropospheric sounding and window channels with a direct sensitivity to solar radiation (also not modelled by the RTM) are obviously biased.

The standard deviation of the observed minus background radiance departures are indicated in figure 1 by the vertical bars (note the length is equal to two times the computed standard deviation). The random departures will be discussed in more detail in section 3c relating to the assignment of observation errors, but overall the values appear qualitatively reasonable from what we know of the instrument noise and the magnitude of typical background errors. The only exception is (again) the shortwave channels where (unexplained) large departures are observed.

In summary, the statistics of observed minus background radiance departures suggest that (with the exception of the shortwave band) there are no gross random or systematic errors in either the data or the RTM. Nevertheless, the observed biases are a significant proportion of the random atmospheric signal we wish the radiance data to correct during the assimilation process. The bias correction strategy is described later in section 3.2.

3. Key elements of the AIRS radiance assimilation system

3.1. The identification of clear AIRS channels

The cloud detection scheme for AIRS is described in McNally and Watts 2003 and the details will not be reproduced here. In summary it is a novel technique for the identification of clear channels at a particular location rather that the more traditional approach of identifying completely clear locations. Departures of the observed spectrum from clear-sky background values are first re-ordered in to a vertically ranked space (i.e. in order of increasing sensitivity to cloud), in which the characteristic signature of cloud becomes monotonic and more readily identifiable. A digital filter is then applied to the departures to reduce the instrument noise (and noise due to errors in the background estimate of atmospheric state). This essentially isolates the pure cloud signal such that the level (or channel in the ranked space) where the cloud contamination first becomes significant can be determined. Channels ranked above this level (i.e. less sensitive to cloud) are retained for assimilation and channels ranked lower (i.e. more sensitive to cloud) are discarded. The result of such an approach is that the data coverage in the analysis is different for each channel. Channels with weighting functions peaking higher in the atmosphere are used extensively whereas channels with lower peaking weighting functions are used significantly less often. The cloud detection scheme has a number of tunable parameters which have initially been set to rather stringent values. This possibly results in the wrongful rejection of some clear data, but ensures that errors due to undetected residual cloud contamination in channels passed as clear are very small (typically less that 0.2K for a mid-tropospheric sounding channel).

3.2. Bias correction

While considerable experience has been obtained over recent years in radiance bias correction for the assimilation of data from low spectral resolution instruments (e.g. AMSU and HIRS see Eyre 1992, Harris and Kelly 2001) the problem for AIRS is considerably more daunting. For this reason, a primary consideration (at least initially) in the design of a bias correction strategy for AIRS was simplicity. Two important characteristics of the biases to be corrected are geographic (or air-mass) dependence and time stability.

The geographic variation of the bias is demonstrated in **figure 2**. It shows the bias evaluated over one month for the tropical region (30N-30S) in the same 4 spectral bands of figure 1. The total length of the superimposed error bar represents (two times) the magnitude of the difference between the tropical bias and the average polar bias (this measure of variability is somewhat arbitrary, but does capture the main type of variability that is observed i.e. north south gradients, see Watts and McNally 2004). In general the pure temperature sounding channels do not show a large variation in bias, except for the channels close to 15 microns. The large bias in these channels has already been attributed to a sensitivity to upper stratospheric / mesospheric temperatures (where the ECMWF model background has large systematic errors). Similarly, the large variability of the bias in these channels is consistent with the strong seasonal and geographic dependence of the upper stratospheric / mesospheric temperature errors. The air-mass dependence of biases in some of the water vapour channels is larger than that for the temperature channels, but is still relatively small compared to the global standard deviation of the signal in these channels. Larger bias variation is also seen in channels with a strong sensitivity to ozone (around 9.5 microns) and some isolated channels with a known sensitivity to un-modelled carbon monoxide (e.g. channel 1867 at 4.58 microns, although the particular asymmetric nature of the bias variation does not have a large value in figure 3). It is interesting to note that, despite the large magnitude, the variability of the short-wave channel biases is very low.

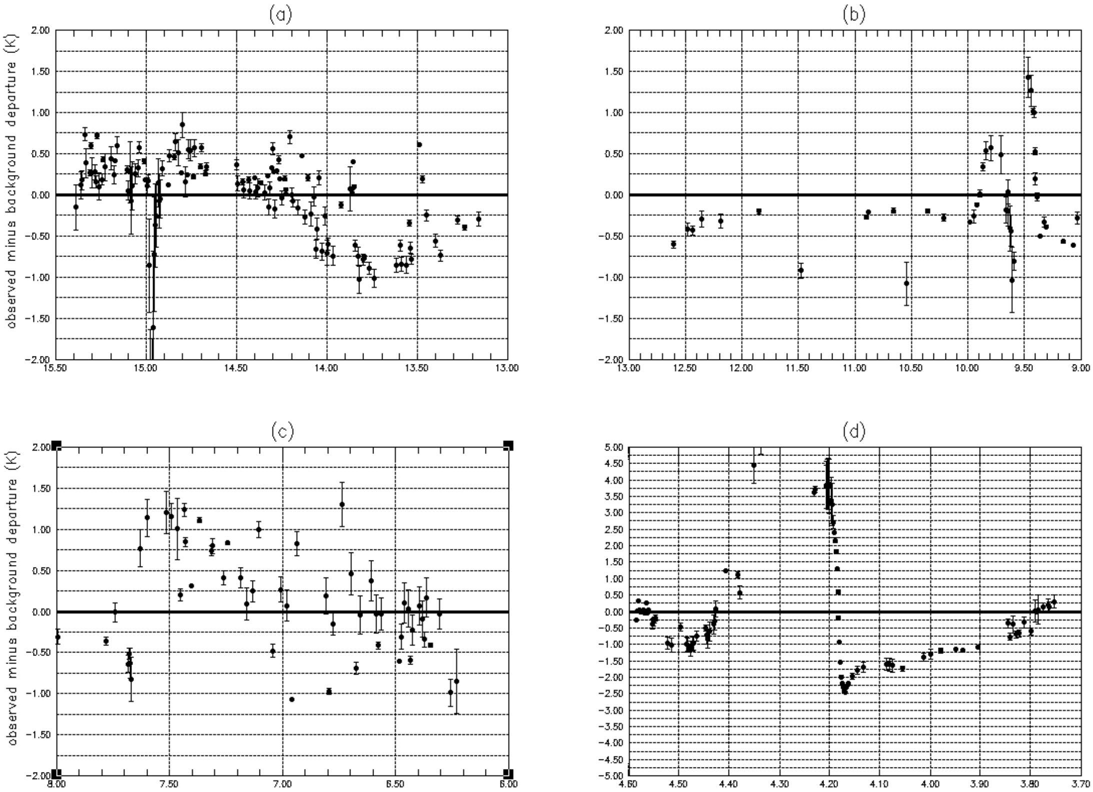

Figure 2 Observed minus background departures evaluated over the tropical region (30N - 30S) for the same 4 spectral bands as figure 1. The black dots show the mean tropical value of the bias and the length of the vertical bar is equal to twice the variation of the bias between the Tropics and Poles. All statistics were evaluated during August 2003 for clear data and the units are degrees Kelvin.

Routine daily monitoring of the AIRS radiance departures suggests that the bias statistics are very stable in time. Again, the only channels that display a significant time variation are those with a sensitivity to the upper stratospheric / mesospheric temperature errors. These show a highly variable bias and illustrate very clearly the seasonal cycle of the known cold bias of the ECMWF model in the winter pole and warm bias in the summer pole.

The dependence of the bias upon instrument scan position is generally very small (other than at the extreme scan edge) and not shown. The only exceptions to this are (again) channels sensitive to the upper stratosphere (where systematic errors in the temperature lapse rate give apparent scan dependent radiance biases) and shortwave channels sensitive to solar radiation.

On the basis of these results (i.e. that the bias in most channels is not highly variable in space or time) an extremely simple bias correction is applied equal to the global mean observed minus background difference in each channel (i.e. the values shown in figure 1). The correction is also fixed in time with no scan dependent component. While this correction is certainly not optimal (as there clearly are some small bias variations which may well be significant) it is considered a reasonable start point for assimilation experiments to begin.

3.3. Observation errors

As with the bias correction an initial aim in assigning observation errors to the AIRS radiances was to keep things as simple as possible and to lean towards conservative (i.e. larger) values. They essentially take one of three different values, giving approximately equal weight to all stratospheric sounding channels (1.0K), equal weight to all tropospheric sounding channels (0.6K) with all other channels (having a significant sensitivity to the surface or water vapour) assigned an error of 2.0K. No attempt will be made to rigorously justify this choice. However, the flat assignment of error is reasonable in that the AIRS instrument noise values do not suggest any dramatic (channel to channel) variations across similar parts of the spectrum. While the instrument noise values in the shortwave are significantly less than for similar sounding channels in the long-wave, statistics of observed minus background radiance departures (shown in figure 1) are similar between the two bands. This suggests that the smaller instrument noise in the shortwave band is being partially offset by larger RTM errors and that a similar assignment of observation error to the long-wave band is justified. For channels sensitive to the surface and water vapour the instrument noise is also very low, but they display large observed minus background radiance departures. While much of this signal may well be due to errors in the background state (and thus should not influence the assignment of observation errors), RTM uncertainties and the effect of undetected residual cloud can be very large in these channels. The assigned observation errors are larger than the standard deviation of observed minus background radiance departures for all channels. The latter represents the summed contribution of the true observation errors (including any errors from the RTM and cloud detection that effectively add to the overall radiance error) and errors in the background estimate of the atmospheric state. As such, they should provide an upper bound for the assigned observation error. However, as inter-channel error correlations and correlations between adjacent soundings are currently ignored, inflation above this upper bound is justifiable.

3.4. Quality control

The analysis is protected from bad data by a simple quality control check based on the departure of the observation from the background. This check takes in to account the error in the observation and the error in the background (mapped in to the space of the observation i.e. radiance in this case). The threshold for rejection is currently set for departures exceeding 3 times the combined standard deviation of observation and background. Considering the variation in values of background and observations error (depending on the particular channel) departures as little as ~3.5K can result in rejection (for the best upper tropospheric sounding channels) and departures as large as ~9K are required to reject some of the water vapour sensitive channels. In practice very few data are rejected by this check, but it must be retained as a safety measure.

3.5. Data selection for assimilation

All channels flagged clear at a given location are assimilated, subject to an additional thinning operation that ensures a minimum horizontal spacing of 120Km (the same as that for other satellite sounding data). The following *a priori* exclusions are applied:

1. Channels beyond AIRS channel 1928 (shorter than 4.46 microns) are excluded pending a better understanding of the large observed biases. This also avoids potential problems of having to selectively use some data only at night.
2. Channels in the ozone band are excluded pending a better understanding of the observed biases, but also to avoid any possible aliasing between ozone signals and temperature adjustments in the analysis.
3. Channels with a significant surface contribution are excluded over land to avoid the possible aliasing of background skin temperature and / or emissivity errors in to atmospheric adjustments.
4. Channels that display obvious air-mass dependent RTM biases are excluded (e.g. previously mentioned channel 1867, sensitive to un-modelled carbon monoxide absorption).
5. All channels measured at the extreme edge of the AIRS scan are excluded.

The configuration selected for this study is by no means optimal, but rather represents a conservative and safe system. Many of the data excluded from this configuration clearly convey valuable information (about e.g. ozone or the land surface), but were considered higher risk options for initial experiments.

4. Assimilation experiments with AIRS

Two different types of experiment have been run to test the impact of AIRS radiances within the ECMWF 4DVAR assimilation system (Rabier et al. 2000). The first is a hypothetical test of the performance of AIRS compared to other (lower spectral resolution) satellite sounding instruments (namely HIRS and AMSUA). In these experiments data from only a single satellite sounding instrument is used in isolation. The second tests the impact of the addition of AIRS to the full operational assimilation, essentially studying what new information is provided over and above the existing global observing system.

4.1. Hypothetical single instrument experiments

Experience in recent years has shown that it is often difficult to demonstrate the impact of a single new instrument on the assimilation system when such a significant volume of satellite data is already present. While in purely objective terms the “value” of a new instrument should indeed be assessed under these conditions (i.e. what information does it bring over and above what we already have), it is also interesting to compare the performance of a new instrument (in isolation) to that of existing sensors. To achieve this a set of three hypothetical experiments have been performed where radiance data from only a single sounding instrument is assimilated. Firstly the AIRS, secondly the AMSUA instrument on NOAA-16 and finally the HIRS instrument on NOAA-16. The NOAA-16 spacecraft is in a similar (but not identical) orbit to the AQUA spacecraft such that the data coverage provided is very similar. There is an AMSUA instrument on the AQUA spacecraft that could have been used for this study, but some channels (6 and 7) cannot be used due to excessive noise and it was considered a fairer test to use an AMSUA with all channels functioning. In addition to the single instrument configurations a baseline assimilation with all satellite sounding radiances removed has been run. This not only aids diagnosis of the results, but provides neutral initial conditions for the other experiments to start from.

Two periods of almost four weeks each have been tested (05/05/2003 to 25/01/2003 and 22/05/2003 to 16/062003) with the baseline system (i.e. all sounding data removed) started 10 days prior to the test periods to avoid any memory contaminating the initial conditions. Apart from the obvious deliberate changes in satellite data usage, the assimilation and forecasting system used is as ECMWF operations of the time (version CY26R3) run at full resolution (i.e. T511 forecast approximately 40Km and T159 analysis approximately 120Km) with a 12 hour 4DVAR window.

The data selection criterion for each instrument (i.e. the choice of which channels are used and which are not) is taken from the ECMWF operational system. In lower-tropospheric temperature sounding channels AMSUA provides by far the best coverage, with significant amounts of AIRS and (to a lesser extent) HIRS infrared data being lost due to the presence of cloud. The more severe loss of data in the case of AIRS (compared to HIRS) is explained by the deliberate stringent tuning of the AIRS cloud detection scheme, but also the reduced horizontal resolution of the available AIRS data (i.e. 1 sounding in 9) reduces the probability of finding a clear location. Note that tropospheric HIRS and AIRS data are always excluded over land. The AMSUA data are not excluded (*a priori*), but in practice very little data are used over land due to difficulties in finding an appropriate surface emissivity value. The data coverage in channels that sense the upper-troposphere is much more similarbetween the three instruments with infrared data only being lost in the presence of very high clouds.

4.1.1. Analysis impact

Analysis increments for temperature in the mid to upper troposphere (model level 36 ~ 400hPa) are displayed in figure 7 (evaluated over the May/June period). **Figure 3a** shows the (root mean square) increments for the baseline assimilation, with significant values restricted to where conventional observations are active. **Figures 3b**,**c,d** shows the increments for each single instrument experiment, with the analysis increments of the baseline (i.e. no radiance) system subtracted. This allows the influence of each satellite instrument upon the analysis to be more clearly seen and not confused by increments due to conventional observations (which are present in all experiments). The largest additional increments due to the radiances are seen in the Southern Hemisphere. In all three cases the single satellite assimilations show larger increments than the baseline over sea (as expected of course), but also show smaller increments than the baseline in many areas where there are conventional data. This is encouraging, demonstrating that the extra work done by the satellites over, for example, the southern Pacific leads to the assimilation system being in better agreement with down-stream conventional observations over South America and Antarctica. The strongest and most widespread excess increments are seen in the AIRS-only assimilation that also shows some of the largest reduction of increments (relative to the baseline) from down-stream conventional observations. A similar signal is seen to a slightly lesser extent with AMSUA radiances and is weakest in the HIRS-only assimilation. The January period (not shown) has slightly stronger increments in the Northern Hemisphere (winter). While the characteristic signal of larger increments than the baseline over sea and smaller increments at down-stream conventional observations is visible (and again strongest for AIRS, weakest for HIRS) it is less clear during this period. The increments (and differences) shown as single level maps in **figure 3** are shown as vertical cross sections in **figure 4**. Apart from the AIRS producing larger increments (away from conventional observations), the AIRS radiances are cause changes to the analysis at a higher vertical scale (than the increments due to either HIRS or AMSUA). Throughout the tropics there are stronger vertical gradients in the increment structure and some vertically localized changes to the temperatures around 400hPa in the band 60 to 40 degrees south.

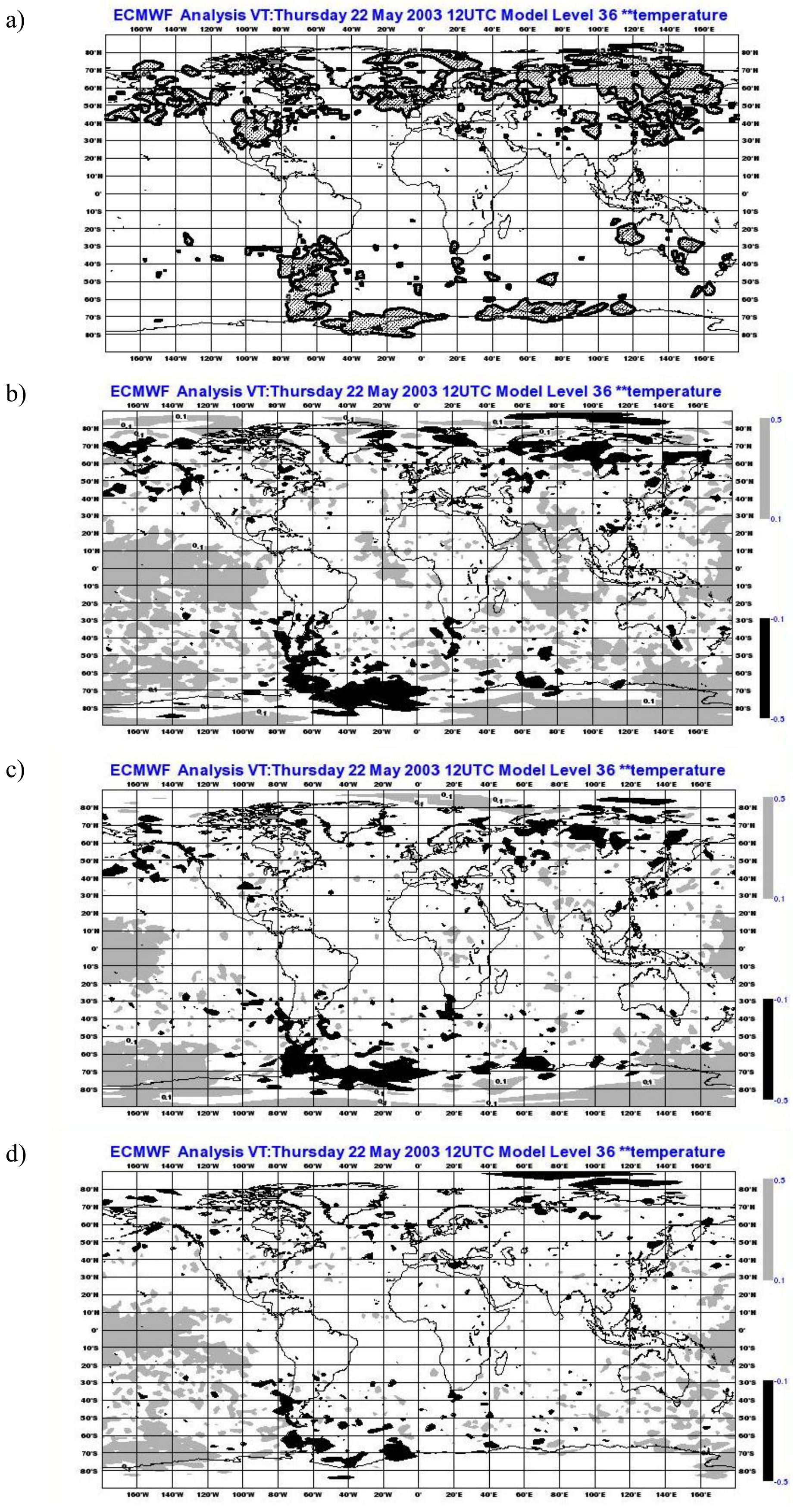

Figure 3 (a) The geographic distribution of root mean square analysis temperature increments at 400hPa for the baseline (no radiance) assimilation (contours begin at 0.5K) during May 2003. The excess increments (i.e. with the values of the baseline system subtracted) for (b) the AIRS only assimilation, (c) the AMSUA only assimilation and (d) the HIRS only assimilation. The units are degrees Kelvin (contours start at 0.1K). Grey indicates increments larger than the baseline, black indicates increments less than the baseline.

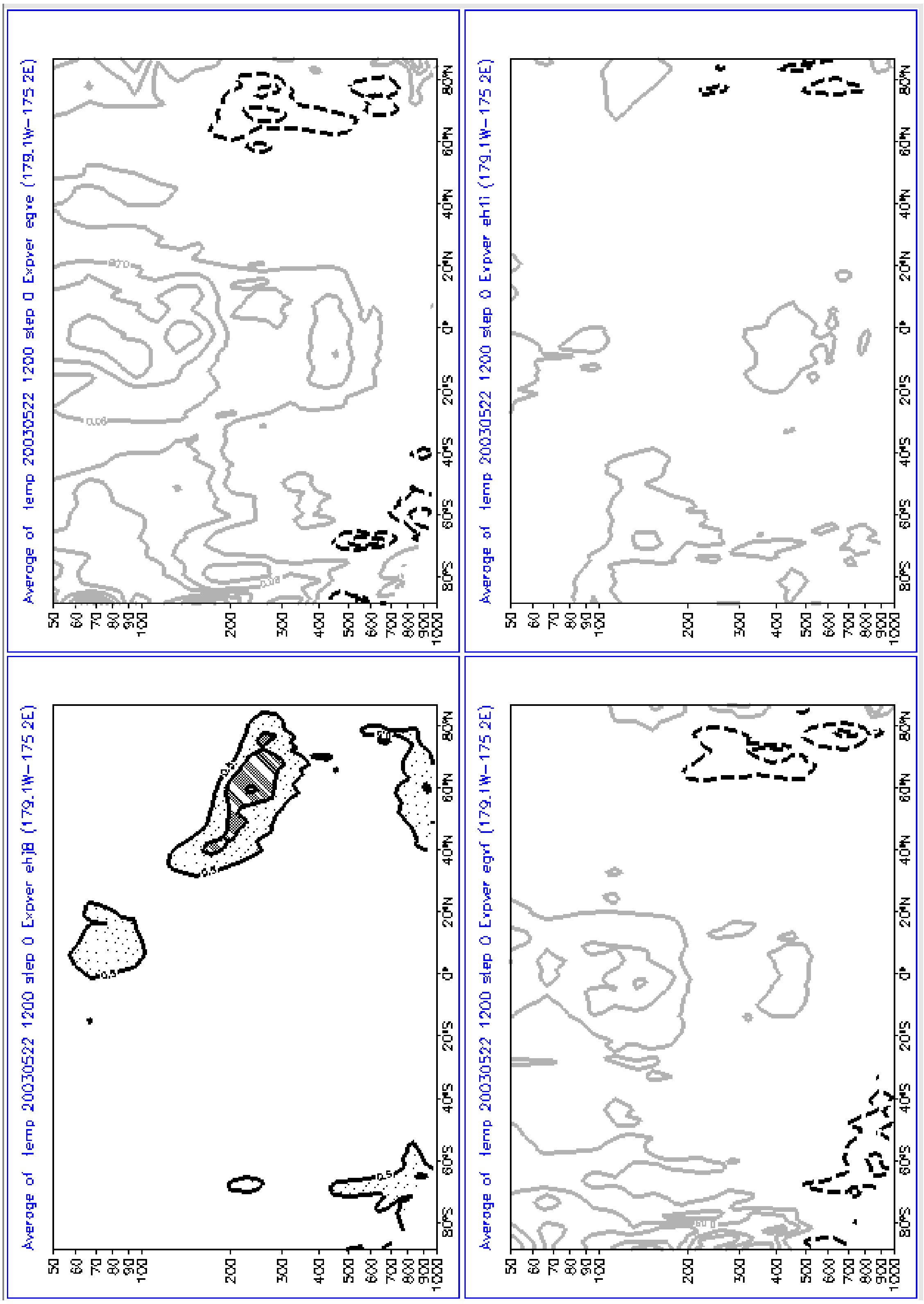

Figure 4 (a) Zonally averaged cross section showing the vertical distribution of root mean square analysis temperature increments for the baseline (no radiance) assimilation (contours begin at 0.5K) during May 2003. The excess increments (i.e. with the values of the baseline system subtracted) for (b) the AIRS only assimilation, (c) the AMSUA only assimilation and (d) the HIRS only assimilation. The units are degrees Kelvin (contours at +/- 0.02 / 0.04 / 0.08 / 1.2 K). Grey indicates increments larger than the baseline, black indicates increments less than the baseline.

4.1.2. Forecast impact

Forecast have been run from the single instrument and baseline assimilations and verified using the ECMWF operational analyses of the time. While this is not particularly clean and may prejudice the results (for example AIRS was not used operationally during these periods and is thus not represented in the verifying analyses) it is probably a reasonable choice and is arguably the most accurate estimate of the atmospheric state globally available. It would certainly be inappropriate to use each experiment's own analyses for verification and the use of radiosonde verification suffers from problems of uneven geographical and temporal sampling.

For simplicity the forecast performance results are shown only in terms of the anomaly correlation of the 500hPa geopotential height, averaged hemispherically. Other parameters (e.g. temperature and wind) and other levels (e.g. 1000hPa and 200hPa) have been examined, but show broadly similar features. Also, as the two periods (summer and winter) also show very similar results, the forecast statistics have been combined together.

In the Northern Hemisphere no single satellite instrument has any clear significant influence upon the forecast quality. This is somewhat surprising given that the assimilation of radiance data did change the increments at some conventional data locations in the Northern Hemisphere. However, it must be remembered that the use of radiance data from all the three instruments is very restricted over land / ice surfaces (as already discussed). It may be that the amount of satellite data being used from a single instrument under these circumstances is simply not sufficient to have a measurable impact on the quality of the mean northern hemispheric forecasts. In the Southern Hemisphere the situation is quite different, with all single instrument assimilations producing forecasts better than those of the no radiance baseline (**figure 5a**). The largest impact is seen from the use of AIRS data, then the AMSUA and finally the HIRS.

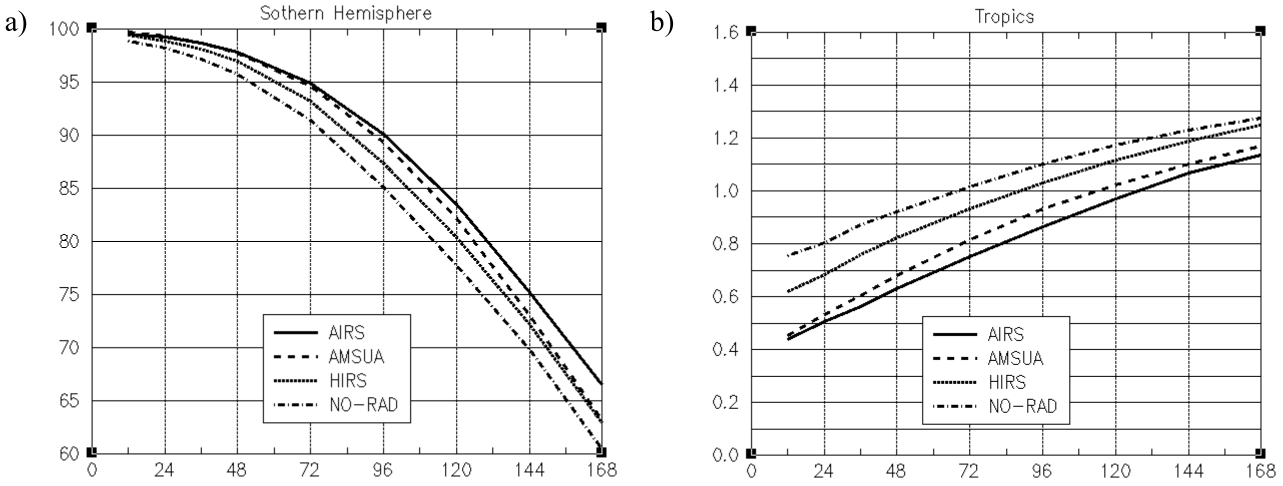

Figure 5 Forecast errors versus forecast range for the single instrument experiments (solid = AIRS, dash = AMSUA, dotted = HIRS and dot-dash = baseline) for (a) mean anomaly correlation of 500hPa geopotential in the Southern Hemisphere and (b) root mean square forecast errors for 200hPa temperature in the Tropics. The scores are averaged over 50 combined cases from both test periods.

The impact upon tropical forecasts is best seen in single level temperature and wind verifications (rather than the geopotential height scores used in the extra tropics). These are generally found to be very similar for all three instruments (all slightly better than the no radiance baseline). The only significant dispersion is seen in forecasts of temperature at 200hPa (**figure 5b**) where the order is the same as that of the Southern Hemisphere.

4.2. Impact of AIRS on the full operational system

Two extended periods have been studied to assess the impact of AIRS radiance assimilation upon the full operational system (i.e. in the presence of all satellite and conventional data). The first period consists of 100 days (2002/12/10 to 2003/03/19) and used version CY25R4 of the ECMWF forecasting / assimilation system . The second period covers 70 days (2004/03/01 to 2004/05/09) and used version CY26R3 of the ECMWF system. In both cases the experiments were performed at full resolution (i.e. T511 forecast approximately 40Km and T159 analysis approximately 120Km) with a 12h 4DVAR window.

The control assimilation against which the AIRS impact is compared (subsequently referred to as "CTRL") uses radiance data from three AMSUA instruments, two HIRS instruments, three geostationary instruments (from GOES and METEOSAT) and three SSM/I instruments as well as wind information from scatterometers, satellite atmospheric motion vectors and all conventional (*in situ*) observations. The experiment (subsequently referred to as "AIRS") is identical except for the additional use of AIRS radiances.

Apart from the obvious different seasonal sampling, the two periods have another (possibly important) difference that should be noted. In the latter period AMSUA data from the NOAA-17 spacecraft was lost (due to an instrument failure) and was replaced by AMSUA data from AQUA.

4.2.1. Changes to the analysis

The increment difference maps (similar to those used in section 4.1 for the single instrument experiments) show a much smaller influence of the AIRS upon the full system (compared to that seen with AIRS upon the no-radiance baseline). In the spring / summer period there are slightly larger increments over the oceans and a small decrease in increments at a number of radiosonde stations (e.g. South America and Antarctica) when the AIRS radiances are assimilated, but the signal is quite weak and rather mixed.

Systematic temperature increments (as opposed to the random increments that have been discussed so far) have been observed in the AIRS assimilation, but in the troposphere these are almost identical to those of the CTRL assimilation during both periods (and are thus not shown). This similarity between the mean increments of the AIRS and CTRL is consistent with the small observed minus background radiance biases seen in tropospheric temperature sounding channels and demonstrates that the simple bias correction scheme applied is effective. However, at the very top of the ECMWF model in the upper stratosphere and mesosphere there are much larger differences in the mean increments of the two systems. The assimilation of AIRS causes a systematic warming of the upper levels at the winter pole and a systematic cooling in the summer pole (and extra-Tropics) and Tropics. This adjustment is associated with known temperature errors in the ECMWF forecast model that have been discussed previously in the context of the observed AIRS radiance biases (section 2). Independent temperature retrievals from the MIPAS instrument on board the ESA ENVISAT satellite agree significantly better with the AIRS adjusted mean analyses at the top of the model (essentially confirming that the AIRS radiance data is correct). Unfortunately the large mean adjustment of the mesosphere has been found to produce smaller, but still significant, oscillating temperature increments at stratospheric levels immediately below (so called "ringing"). This effect has been diagnosed in simulation and traced to a weakness in the 4DVAR background error formulation (Rabier and McNally 1998), the statistics not representing such large correlated background errors. During the second test period the uppermost AIRS sounding channels (which drive the temperature adjustment at the top of the model and the associated ringing below) were removed from the assimilation pending improvements to the background formulation or correction of the model error by other means (e.g. improved physics). Offline experiments showed that the removal of the upper stratospheric channels had no measurable impact upon the tropospheric analysis or forecast quality.

Despite AIRS having many channels sensitive to water vapour, no significant changes in the analyzed humidity have been observed in either of the two periods tested. There are two reasons for this. Firstly, relative to the temperature sounding channels, the humidity channels have been assigned high observation errors of 2K. This inflation was a precautionary measure discussed in section 3.3. The deliberate down-weighting of the data coupled with the fact that the ECMWF humidity analysis is already highly constrained by existing satellite observing systems (such as the SSM/I controlling the boundary layer and a combination of HIRS and GEO radiances controlling the upper tropospheric humidity) mean that large humidity changes are not expected.

4.2.2. Forecast impact

Forecasts have been run from the analyses that assimilated AIRS radiances and compared to those from the CTRL system. For the first period the verifying analysis is ECMWF operations that used no AIRS data (a choice that may slightly penalize the AIRS system). For the second period the verifying analysis is again ECMWF operations, but at this time AIRS radiance assimilation had become part of the operational system (and thus this choice of verification may slightly penalize the CTRL system). However, it should be noted that, for all but very short-range forecasts, the choice of verification in this context is not expected to have any measurable effect.

Figure 6 shows a comparison of hemispheric averaged mean forecast scores for 500hPa geopotential height (the same measure of forecast quality used for the single instrument experiments). In both periods and in both hemispheres there is a very small, but consistent improvement in forecast skill at all ranges when the AIRS data are assimilated. The only noticeable change between the two periods is that the improvement due to AIRS in the Southern Hemisphere is slightly larger for the second period. Other levels (e.g. 1000hPa and 200hPa) and other parameters such as wind have been verified. The results (which are not shown) are very similar. The statistical significance of these changes has been examined using a *t-test*. It has been found that at day 3 and day 5, the differences between the forecast errors of the AIRS and CTRL are statistically significant (to better than 1% in the Northern Hemisphere and better than 5% in the Southern Hemisphere), but at day 7 the changes are not significant (to better than 10%).

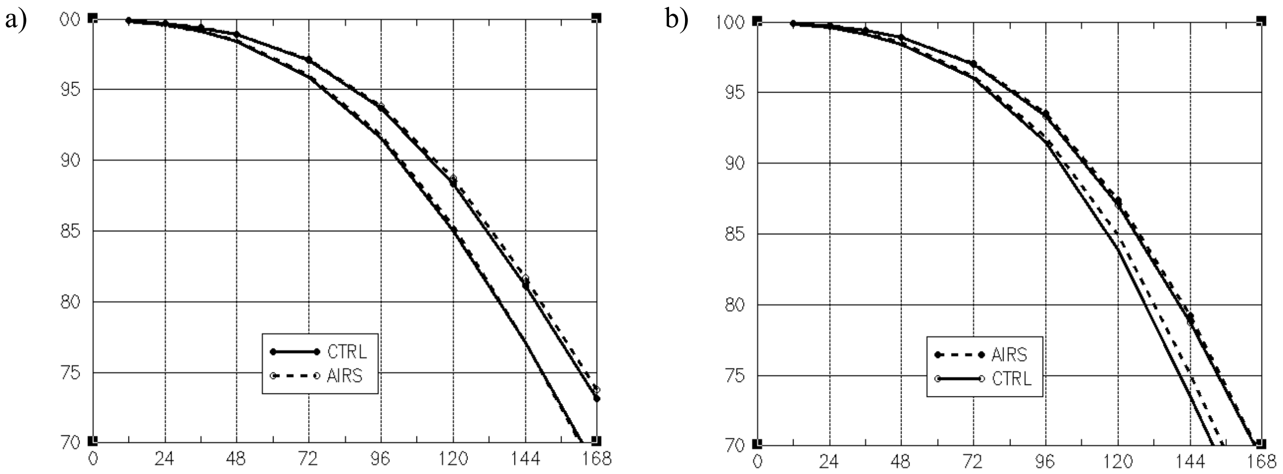

Figure 6 Forecast errors versus forecast range in the AIRS assimilation (dash) and CTRL assimilation (solid). Mean anomaly correlation of 500hPa geopotential in the Northern Hemisphere (top two curves) and Southern Hemisphere (bottom two curves) during (a) the 100 cases of the winter period and (b) 70 cases of the spring / summer period.

The impact of AIRS radiances upon tropical forecasts of single level temperature and wind is generally neutral and not shown. The only exception is that (as with the single instrument experiments) the use of AIRS improves forecasts of 200hPa temperature (see **figure 7**). Note that the root mean square forecast error is initially greater with AIRS in the statistics of the first period, but then crosses to become better than the CTRL further into the forecast range. This reflects the use of a verifying analysis (i.e. ECMWF operations) that did not use AIRS radiances (which is not the case for the second period and no crossing of the curves is observed).

The geographical distribution of forecast errors has been examined (the maps are not shown). It appears that the improvements gained with AIRS are not localized to any particular area, being spread fairly evenly over the globe. While there is a net gain using AIRS in the hemispheric average scores, the maps clearly show that the situation is a mixture of good and bad forecasts (particularly at day 7). Indeed during the first period there is an almost even mix of good and bad forecasts at day 7 in the Southern Hemisphere (explaining the observed neutral impact of AIRS on the hemispheric score).

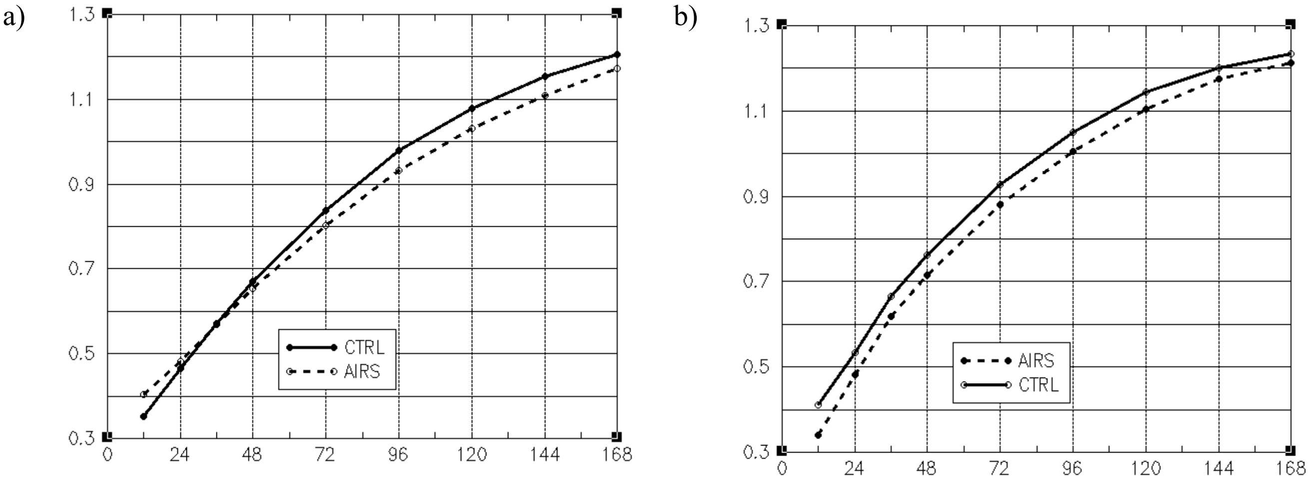

Figure 7 Root mean square forecast errors for 200hPa temperature in the Tropics for the AIRS assimilation (dash) and CTRL assimilation (solid) during (a) the 100 cases of the winter period and (b) 70 cases of the spring / summer period.

4.3. Discussion of results

In the hypothetical single instrument experiments it was found that the use of radiances from any of the three instruments improved the analysis fit to conventional data (relative to the no radiance baseline system). However, there is a clear order, with the AIRS demonstrating the strongest impact upon the analysis (having the largest increments over the oceans and smallest increments near in situ data), followed by the AMSUA and the smallest changes caused by HIRS. In addition, the increments due to AIRS were seen to have a higher vertical scale than either of the other two instruments. The same order is found in the quality of forecasts from each of the three single instrument experiments. In the Southern Hemisphere all instruments improved over the no-radiance baseline, with the AIRS performing best, followed by the AMSUA and finally the HIRS. While it is difficult to make a rigorous causual link between the size and structure of the observed analysis increments and the subsequent forecast performance, it is unlikely to be a coincidence that the same order in observed in both. The results suggest that, despite losing significant amounts of data in the lower troposphere due to cloud contamination, the assimilation of AIRS radiances constrains important atmospheric structures elsewhere and reduces the development of forecast errors (to a greater extent than either AMSUA or HIRS).

The results of the single instrument experiments are very encouraging, but their importance should not be overstated or extrapolated beyond the very hypothetical situation of only having one satellite sounding instrument available. Under these artificial conditions errors in the background (that we are looking to correct using the radiance data) may be larger and distributed differently (e.g. in the vertical) compared to a realistic assimilation environment (where the errors are more highly constrained by other satellite data). Indeed the results of assimilation experiments with the full ECMWF system show a much more modest impact of AIRS. The analysis increments do show a better fit to conventional data in some areas of the Southern Hemisphere (at least for the second period) and forecast errors are reduced when AIRS radiances are used, but the magnitude of the improvement is small and warrants some discussion.

The AIRS data comes from only one platform and the data available for assimilation is highly sampled (only 0.15 percent of the radiances that are measured are actually available in near-real-time). If we couple this with the deliberate conservative use of the data inside the 4DVAR assimilation scheme we possibly should not expect a dramatic impact from this initial introduction of AIRS (in a system already densely populated with other satellite data). However, large improvements may also be limited by the quality of the CTRL system, which, averaged over the two (substantial) periods tested, is arguably very high. Apart from the average performance being very good, a detailed time series analysis of forecast skill shows that the CTRL system produces very few poor forecasts or "busts". For example, during the 100 day trial **no** day-5 forecasts of 500hPa height scored less that 60% anomaly correlation in either the Northern or Southern Hemisphere. Verified over the much smaller European area, still only 6 forecasts from the CTRL could be considered as busts (scoring worse than 60% at day 5). In four of these cases adjoint sensitivity diagnostics (Isaksen 2004) trace the forecast bust to problems in the initial conditions (in the other two perturbations to the initial conditions had no effect), but only one of these (24 Feb 2003) was sufficiently clear in the important (sensitive) areas for any tropospheric AIRS radiances to be used. In this one case that was relatively free from high clouds it appeared that the some of the analysis increments due to AIRS did correlate with the adjoint sensitivity perturbations, but many did not and AIRS did not fix the forecast bust. Overall it is difficult to argue from the evidence of these trials that the assimilation of AIRS is fixing bad forecasts on any regular basis. It appears more that the assimilation of AIRS (with the current configuration) is having a small, but relatively consistent positive impact upon the mean forecast skill. It should be noted that the mean forecast skill of the second period is lower and we observe a slightly larger impact of AIRS. While this may be simply due to different sampling, we may speculate that the loss of the fully functioning NOAA-17 AMSUA instrument in a separated morning orbit may not be fully compensated by the use of AQUA AMSUA in a redundant pm orbit (with two channels unusable due to noise).

5. Conclusion and future development

The key elements of the AIRS assimilation have been described. All have been constructed with simplicity in mind and tuned towards a conservative (i.e. cautious) use of the radiance data. However, even with such a conservative configuration, the use of AIRS radiances has shown some very encouraging results. The AIRS clearly out-performs the two lower spectral resolution instruments (AMSUA and HIRS) in single instrument experiments and has shown a modest, but positive impact upon the full operational assimilation and forecasting system (averaged over a substantial sample of 170 cases). The assimilation of AIRS radiances was implemented operationally at ECMWF in September 2003.

There are, of course, areas where we have considerable scope to improve, namely by a better use of the data we already have in the system and extending the assimilation to exploit more of the radiance information measured by the AIRS instrument. These are described below in no particular order.

The understanding and correction of systematic errors (mainly from the RTM) is likely to remain a difficult task. However, there are encouraging indications (e.g. Strowe 2003) that the very detailed spectral measurements provided by AIRS are allowing considerable advances to be made in diagnosing individual sources of the error (i.e. within our spectroscopic knowledge and line-by-line parameterization) which should ultimately lead to an improved RTM. In the meantime we will have to improve the treatment of radiance biases within the assimilation system. The current flat bias has proved to be adequate so far, but its continued use is likely to hinder efforts to make more aggressive use of the AIRS data (e.g. the reduction of observations errors and modelling of inter-channel correlations). A more sophisticated (physically based) air-mass dependent bias correction is described in Watts and McNally 2004, and preliminary results show a useful reduction of systematic errors, particularly at high (polar) latitudes.

Much of the AIRS information from the short-wave part of the spectrum has not been used in this study, but it clearly contains valuable information that could compliment the use of long-wave channels (the shortwave channels also have sharper weighting functions). While we are probably far from being able to use short-wave radiances contaminated by solar radiation, significantly more channels than are currently used could be assimilated at night. This will require careful treatment of the larger systematic errors observed in the shortwave and a more sophisticated description of surface emissivity (over land and sea) than the constant value used so far.

If we wish to realize the full potential of AIRS to constrain small vertical scale features we will have to increase the number of channels used in the analysis. A possibility for communicating more of the AIRS spectrum to the analysis in an efficient way is the use of principal components (or eigenvectors). It has been shown by a number of authors (e.g. Huang and Antonelli 2001 and Goldberg et al. 2003) that a given AIRS spectrum can be efficiently represented in terms of a truncated set of eigenvectors (typically only 200 eigenvectors out of the original 2378). With an appropriate normalization, the truncation can be chosen to selectively remove eigenvectors that describe noise while retaining those that convey the atmospheric information. A subset of radiances reconstructed from the truncated set of eigenvectors are essentially de-noised by this process and their assimilation incorporates information from a much larger number of channels in an efficient way. While our understanding of this approach is far from complete it is a high priority for further investigation.

Improvements in out treatment of clouds are as likely as any of the above to increase the impact of AIRS. The current approach of identifying clear channels is successful, with a number of refinements to the algorithm (currently under development) and access to the full spatial resolution data expected to increase the volume of clear AIRS data that can be assimilated. However, more substantial progress may only come from the use of cloud contaminated radiances in the analysis. The fundamental difficulty of separating the cloud signal from the atmospheric signal (which is typically only a few tenths of a Kelvin) may prove too difficult for highly contaminated radiances (e.g. where the cloud signal is many tens of Kelvin). However, the treatment of lower level contamination is a more tractable problem and will be pursued.

References

Aumann, H. H., M. T. Chahine, C. Gautier, M. D. Goldberg, E. Kalnay, L. M. McMillin, H. Revercomb, P. W. Rosenkranz, W. L. Smith, D. H. Staelin, L. L. Strow, and J. Susskind, 2003: AIRS/AMSU/HSB on the Aqua Mission: Design, Science Objectives, Data Products, and Processing Systems. *IEEE Transactions on Geoscience and Remote Sensing*, **41**, 253-264

Fourrie and Rabier 2002 The use of advanced sounders in cloudy conditions. *Proc. 12th International TOVS study conference*, Lorne Australia February 2002.

Goldberg, M. D., Y. Qu, L. M. McMillin, W. Wolf, L. Zhou, and M. Divakarla, 2003: AIRS Near-Real-Time Products and Algorithms in Support of Numerical Weather Prediction. *IEEE Transactions on Geoscience and Remote Sensing*, **41**, 379-389.

Isaksen, L., M. Fisher and E. Andersson: The structure and realism of sensitivity perturbations and their interpretation as 'Key Analysis Errors'. *ECMWF 2004 Technical Memorandum* 445 (Available on http://www.ecmwf.int/publications/)

Matricardi M., F. Chevallier and S. Tjemkes 2001 An improved general fast radiative transfer model for the assimilation of radiance observations. *ECMWF Technical Memorandum* 345.
(Available on http://www.ecmwf.int/publications/)

McNally A.P. 2002 A note on the occurrence of cloud in meteorologically sensitive areas and the implications for advanced infrared sounders *Q. J. R. Meteorol. Soc.* 128, 2551-2556

McNally, A.P. and P.D. Watts A cloud detection algorithm for high-spectral-resolution infrared sounders *Quart. J. Roy. Meteor. Soc.* 2003 129 pp 2411-3423

Prunet P., Thepaut J.N., and Casse V., 1998 The information content of clear sky IASI radiances and their potential for numerical weather prediction. *Q. J. R. Meteorol. Soc.*, 124, 211-241.

Rabier, F. and McNally, A.P. 1993 Evaluation of forecast error covariance matrices. *ECMWF Technical Memorandum* 195. (Available on http://www.ecmwf.int/publications/)

Rabier F., McNally A.P. 1998 Evaluation of forecast error covariance matrices *ECMWF Technical Memorandum* 195. (Available on http://www.ecmwf.int/publications/)

Rabier, F. and Järvinen, H. and Klinker, E. and Mahfouf, J-F. and Simmons, A. 2000: The ECMWF operational implementation of four-dimensional variational assimilation. II: Experimental results with simplified physics. *Quart. J. Roy. Meteor. Soc.*, **126** , 1143-1170.

Strow, L. L., S. E. Hannon, S. De Souza-Machado, and H. E. Motteler, 2003: Validation of the AIRS radiative transfer algorithm using ECMWF data fields. *Proc. SPIE Remote Sensing of the Clouds and Atmosphere* VII, Crete, Greece, 4882, 90-99.