

The assimilation of AIRS radiance data at ECMWF

A.P. McNally, P.D. Watts, J.A. Smith
*European Centre for Medium-range Weather Forecasts
Reading, RG2-9AX, UK*

Introduction

The NASA AQUA spacecraft was launched in May 2002 and a subset of radiance data from the AIRS instrument (324 channels from 2378 at 1 sounding location out of 18) have been made available to ECMWF in near-real-time (NRT) since the end of October 2002. Prior to this date a significant amount of technical development was achieved using simulated AIRS data sets routinely provided by NOAA/NESDIS and some early scientific validation was also performed using a pre-released “focus-day” of data (for 24 August 2002). As a result of this preparation the processing (i.e. cloud-screening and monitoring) and assimilation impact experiments were able to commence almost immediately following the arrival of the real AIRS data. This memo documents the progress to date and concludes that we now have a safe “conservative” assimilation system for AIRS which should be considered for a day-1 operational implementation.

Initial evaluation of the AIRS radiance data

Before any assimilation experiments were performed the degree of consistency between AIRS radiances simulated by our radiative transfer model (RTM) and the radiances measured by the instrument had to be evaluated. This not only quantifies the accuracy of the instrument spectral characterization, but also the accuracy of the RTM. Figure 1 shows mean observed minus background radiance departures for data determined as clear (the identification of clear data will be discussed in the next section). It can be seen that the bias indicated by the red line (corresponding to the post-launch spectral response functions) is generally quite small and certainly at a level where it becomes difficult to attribute the source of the bias (i.e. systematic errors in the instrument, RTM or the background fields of temperature and humidity). The largest biases are found in channels with a strong sensitivity to water vapour and the shortwave channels with a sensitivity to solar radiation (not currently modelled by our RTM).

The departure statistics have been found to be very stable in time, but also display very little air-mass dependence. Figure 2 shows zonally averaged biases for channels in the 15 micron band of AIRS ranked vertically from the surface to the top of the atmosphere. On the right hand axis the pressure at which the channel Jacobian reaches a maximum is shown (i.e. its peak sensitivity). For channels sensitive to the mid-troposphere and lower stratosphere there is only a small geographical variation in the bias. Channels peaking near the top of the model display much larger variations with latitude, but these variations are consistent with our knowledge of model systematic temperature errors (from AMSUA and HALOE) and are unlikely to be due to air-mass dependent biases in the AIRS data. Window channels peaking at the surface also show larger systematic variations, but it is impossible to distinguish if these are true biases in the AIRS data or systematic errors in our modelling of the surface emission (i.e. from the model skin temperature and emissivity) or indeed problems detecting cloud over these surfaces. On balance it was concluded that air-mass dependent biases were not significant and that

assimilation experiments could commence with a very simple static (in time) and flat (in space) bias correction applied to the AIRS radiance data. Standard deviations of observed minus background departures (not shown) were found to vary between less than 0.5K for the best channels (dry temperature sounding channels) and 1-2K for the worst channels (i.e. those sensitive to water vapour and temperatures around the surface / stratopause). In most cases the random departure statistics for AIRS were found to be very consistent with values for similar channels on other sensors (e.g. HIRS / AMSUA) that have been observed for many years.

The identification of clear AIRS channels

The cloud detection scheme for AIRS is described in *McNally and Watts 2003* and will not be reproduced here. In summary it is a novel technique for the identification of clear channels at a particular location rather than the more traditional approach of identifying clear locations. Figure 3 shows the location of clear data (shown by red) in 3 different channels for a typical case. The coverage of clear data in the sounding channel sensitive to the upper troposphere / lower stratosphere is only occasionally interrupted by very high cloud, but significantly more data is lost in the channel sensitive to the mid-troposphere. The window channel coverage is sparse and corresponds to the limited number of locations where the atmosphere can be identified as completely clear to the surface. The cloud detection scheme has a number of tuneable 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 flagged clear are very small.

Assimilation configuration for AIRS radiances in 4DVAR

Following a reasonably comprehensive set of experiments carried out over a 1 month initial trial period in October / November 2002 (at a reduced T159 resolution and using a 6hr 4DVAR) we converged upon an initial configuration for the use of the AIRS radiances, the key elements of which are summarized below:

- Input radiance data consists of sampled 1/18 locations and 324/2378 channels
- No assimilation of channels in the O₃ or 4.2 micron band (approximately 100 of the 324 channels)
- Over sea, all channels flagged clear (including window channels) are assimilated
- Over land, only long-wave channels peaking above approximately 400hPa are assimilated
- Soundings are thinned to a horizontal spacing of 120Km preferentially retaining the clearest
- Flat (single global number rather than varying) bias correction used for each channel
- Simple observation error assigned to different blocks of channels (0.6K for dry tropospheric temperature channels away from the surface and stratopause, 1.0K for

stratospheric temperature sounding channels and 2.0K window channels and channels sensitive to water vapour).

The testing that led to this configuration considered a variety of performance measures including the size of analysis increments, the fit to other observations and forecast impact. The configuration selected is by no means optimal, but rather represents a reasonably safe *baseline* system that could be tested as a candidate for day-1 operational implementation. Many of the data excluded from this configuration clearly convey valuable information, but were considered higher risk options until further work on elements such as cloud detection, modelling land surface emission, ozone and solar radiation could be done. The observation errors are similarly set to conservative levels. While these almost certainly overestimate the true observation errors, we currently take no account of inter-channel error correlations and thus some degree of inflation is justified.

Impact experiments using AIRS radiance data

The baseline AIRS configuration described above has been tested at full resolution in 12hr 4DVAR using cycle 25R4 of the IFS between 10 Dec 2002 and 19 March 2003 (a total of 100 cases) and is subsequently referred to as “AIRS”. The control against which the AIRS impact is compared (subsequently referred to as “CTRL”) is generally the operational system, but a research department experiment was used prior to cycle 25R4 being implemented in operations.

Changes to the analysis

Figure 4 shows a difference map (AIRS minus CTRL) of RMS analysis temperature increments at 500hPa (averaged over a ten day period in December 2002). While the contour interval is extremely fine (shading starting at 0.1K) the map shows that there are slightly larger increments over the oceans (where most of the AIRS radiances are used) and a small (but fairly consistent) decrease in increments at radiosonde stations when the AIRS radiances are assimilated (the large increase over central Africa originates from the use of AIRS data over lake Chad that is treated as “sea” in the assimilation). The reduced increments at radiosonde stations is an encouraging diagnostic and shows that the extra work being done by the AIRS data in the analysis improves the agreement with radiosonde data.

Systematic analysis increments in temperature and humidity for the AIRS and CTRL assimilations are shown in figures 5 and 6 respectively as zonally averaged cross-sections (evaluated over the same 10 day period). It can be seen that the mean increments are generally rather small and to first order are very similar in the troposphere and lower stratosphere between the two systems (with the AIRS increments possibly being slightly larger). In the upper stratosphere there are much larger differences in the geographical and vertical distribution of mean increments. As yet these have not been investigated fully, but may reflect the ability of the AIRS channels to resolve finer vertical structures near the stratopause (we hope to use lidar and MIPAS data to examine this issue further). The similarity (in the troposphere) between the mean increments of the AIRS and CTRL is consistent with the small observed minus background radiance biases discussed previously.

Generally the small changes to the analysis illustrated above are not large enough to significantly impact the usual observation fit statistics computed over large areas

(except in the temperature fits above 10hPa where the AIRS slightly improves over the CTRL). However, the fact that the assimilation of AIRS does not change the overall fit to other observations (temperature or wind) is also very encouraging as it suggests that using a very simple “flat” bias correction applied to the AIRS radiances has no detrimental effect.

Forecast impact

Forecasts have been run from the analyses that assimilated AIRS radiances and compared to those from the CTRL system. Figure 7 shows forecast error difference maps (AIRS minus CTRL) for 500hPa height, each system verified using its own analyses and averaged over the first 50 cases of the trial. Blue shading indicates where the AIRS forecasts are better and yellow where they are worse than the CTRL. It can be seen that the assimilation of AIRS has reduced forecast errors at all ranges. The impact is first seen in the short-range (day-3) forecast of the Southern Hemisphere and then in the day-5 forecasts of both hemispheres. The positive signal increases with increased forecast range for the Northern Hemisphere, but becomes marginal beyond day-7 in the Southern Hemisphere (where the situation appears to be more a mixture of good and bad forecasts). No signal is seen in the tropics, but this is due to the choice of forecast variable (i.e. Z500). A positive impact of the AIRS upon the forecasts of tropical temperatures (not shown here). Larger samples (100 cases) of area-averaged mean forecast scores for 500hPa geopotential height have been generated. However, it should be noted that these have been computed using the operational / CTRL analyses for verification, a choice that may slightly penalize the AIRS system. Over 100 cases there is still a very small, but very consistent improvement at all ranges in the Northern Hemisphere (the results of significance testing are contained in tables 1 and 2 show that the improvement is statistically significant at the 1% level for day-5). For the European area (imbedded in the Northern Hemisphere statistics) the positive impact is marginally clearer, but less significant. In the Southern Hemisphere, only a slight improvement is seen at day-3 (significant at the 5% level) and beyond this no improvement is seen over the CTRL (the negative impact at day-10 was not found to be significant < 10%). The verification of temperature forecasts from the 2 systems is generally consistent with the height results in the mid-latitudes, but they additionally show a positive impact of the AIRS in the tropical temperatures at 200hPa. The same statistic for the southern hemisphere shows larger RMS errors when AIRS data are used, but a closer investigation indicates a large systematic difference between the AIRS and CTRL analyses, localized to the edge of the Antarctic continent and not evident at any other level than 200hPa.

Wind forecasts from both systems have been verified and scores at 1000hPa (typical of other levels) display similar signals to the height verifications, although a positive impact due to AIRS in the Southern Hemisphere is slightly more evident. In the statistical significance testing of the forecast impact (shown below) red indicates a positive impact due to AIRS and blue a negative impact. The percentage figure indicates the level at which a *t*-test found the results statistically significant. If no significance better than 10% is found the result is marked with an X

Forecast Range	Northern Hemisphere	Southern Hemisphere	Europe
day-3	5% / 1%	5% / 10%	X / 2%
day-5	0.1% / 1%	10% / X	10% / 5%
day-7	X / X	X / X	X / X

Table 1 Significance testing of 1000hPa (first figure) 500hPa (second figure) height forecast verifications

Forecast Range	Northern Hemisphere	Southern Hemisphere	Europe
day-3	X / 5%	0.1% / 0.1%	10% / 0.5%
day-5	0.1% / 0.1%	2% / 5%	5% / X
day-7	0.1 / 2%	X / X	X / 10%

Table 2 Significance testing of 1000hPa (first figure) 500hPa (second figure) wind forecast verifications

Discussion of results

The assimilation of AIRS radiances with the baseline system described here shows no adverse effects in the analysis (in terms of the fit to other observations) and slightly reduced analysis increments at radiosonde locations. Overall the forecast performance of the baseline AIRS assimilation scheme is encouraging, essentially showing a consistent positive impact in most areas and parameters. However, averaged over the 100 cases the impact is small and warrants some discussion. The assimilation configuration is clearly conservative and a variety of further enhancements (many of which are already being tested) will be described in the next section. However, large improvements over the CTRL may also be limited by the quality of the CTRL system itself. The average level of forecast skill for the CTRL (that currently uses radiances from 3 AMSUA, 2 HIRS, 3 GEOS and 3 SSM/I instruments) is very high and over the period tested was significantly better than forecasts from any other NWP centre. Furthermore, a time series analysis of forecast skill shows that the CTRL system produces very few poor forecasts or “busts”. During the 100 day trial **no** day-5 forecasts of 500hPa height scored less than 60% anomaly correlation averaged over either of the hemispheres. Verified over the much smaller European area, still only 6 day-5 forecasts from the CTRL scored worse than 60%. In 4 of these cases the AIRS system improved the forecast by 10% or more (4 AIRS forecasts scored worse than 60% over the period, but the CTRL was never 10% better). Most of the cases where AIRS improves the poor forecasts correlate with when adjoint sensitivity perturbations to the initial conditions (rather than “forced” perturbations) were found to have a large effect. However, the improvements are far less dramatic than those achieved (retrospectively) by the sensitivity perturbations. Usually cloud was found to obscure many of the sensitive locations (resulting in very few tropospheric AIRS radiances being used). In the one case that was relatively clear (24 Feb 2003) it appeared that the some of the analysis increments due to AIRS did correlate with the sensitivity perturbations, but

many did not. Overall it is difficult to argue that the assimilation of AIRS is dramatically 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 forecast skill.

Work in progress and next steps

On the basis of the results presented here, AIRS radiances were introduced into the ECMWF operational assimilation system on the 7th October 2003.

The areas where we feel the AIRS system can be enhanced in the short-term relative to the baseline configuration are listed here in no particular order:

Improvements to the cloud detection will remain a priority. We are reasonably confident that the current system is safe, but we know clear data is being discarded. Developments well under way by P Watts and J Smith aim to exploit collocated visible data and cross-spectral checks to allow a more skilful identification of clear channels.

The decision to exclude tropospheric channels over land was made early in the development process and was mainly related to the occurrence of large increments thought to be associated with missed cloud. Since then many improvements have been incorporated to the cloud detection and the use of tropospheric radiances over land will be reviewed.

The issue of bias correction has partially been investigated. While all the results presented here have used a flat bias correction for AIRS radiances, parallel experiments have been performed using the traditional air-mass dependent scheme extended to AIRS. The air-mass dependent correction slightly degrades the fit to radiosonde temperature data at 200hPa and 70hPa relative to the flat correction (suggesting some of the air-mass dependence is systematic model error), but slightly improves the temperature fit at 100hPa. For the 45 cases tested, assimilating AIRS with the air-mass dependent bias correction results in a slightly degraded performance in forecasts of Northern Hemisphere 500hPa height. Thus from the results so far there is no strong incentive to employ an air-mass dependent bias correction for AIRS. G Kelly is currently investigating to what extent the air-mass dependent corrections applied to other sensors (e.g. AMSUA and HIRS) could be relaxed back to a simple flat correction. If this could be achieved without any loss of NWP performance it may assist our parallel efforts to extract CO₂ information.

Much of the AIRS information from the short-wave part of the spectrum has not been used. 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. However, there are some important additional issues related to the accuracy of the RTTOV model around 4 microns (for channels located on the steepest part of the absorption band) and the RTM may need to be upgraded to 101 vertical levels (currently it has only 43) before significant progress can be made.

Finally the observation errors assigned to the AIRS radiances will be reviewed in line with improvements to the cloud detection, bias correction and our general understanding of the characteristics of the AIRS data.

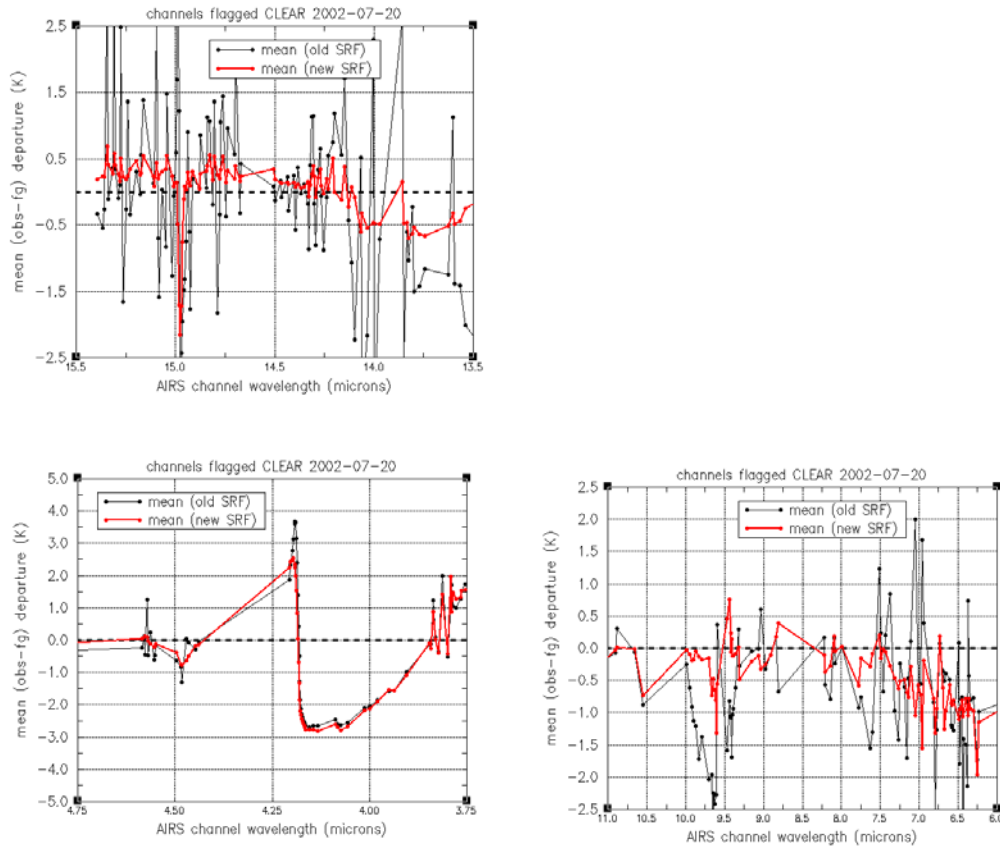


Figure 1 Mean observed minus background radiance biases for clear data. The red line corresponds to the RTM using spectral response functions (SRF) evaluated after launch, the black to pre-launch SRF.

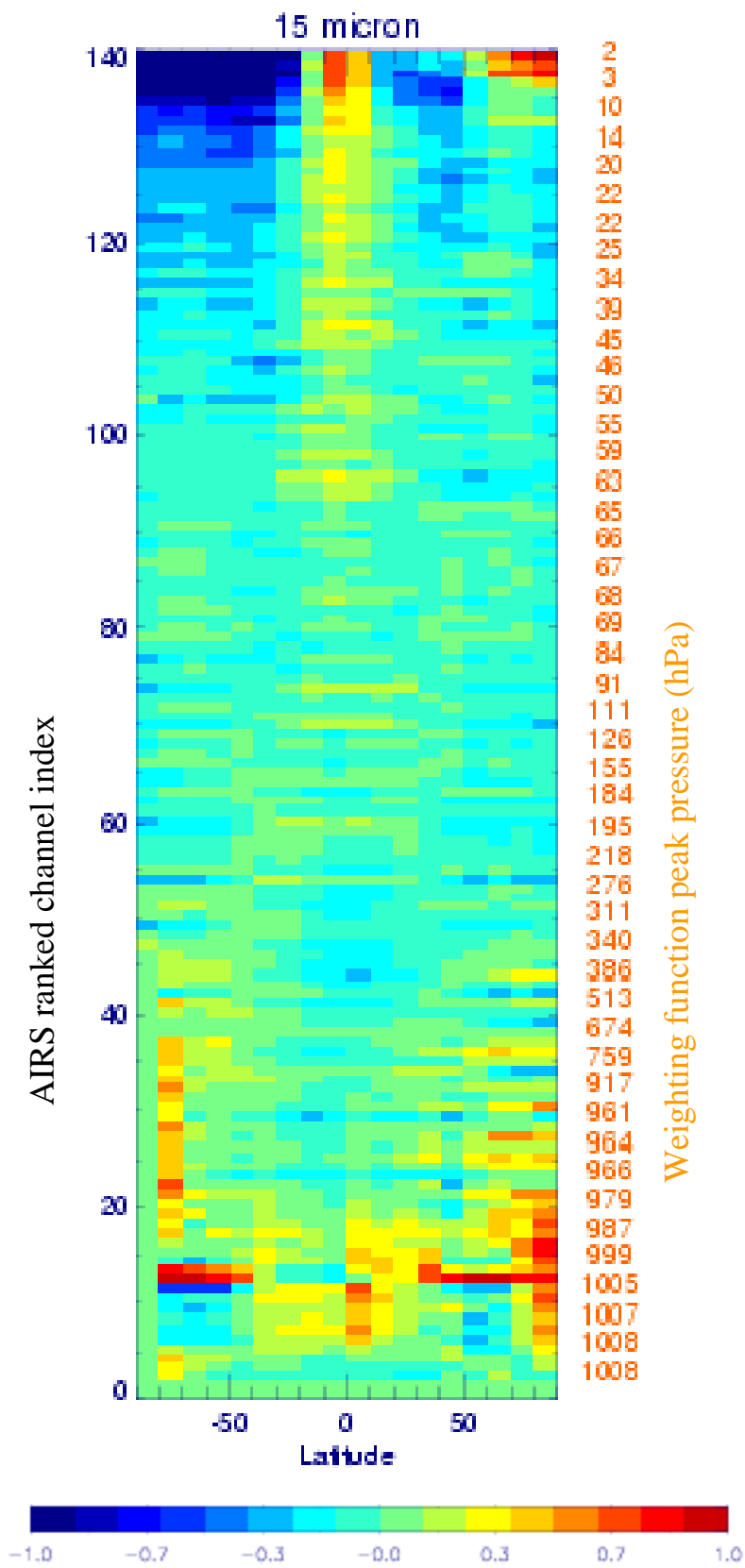


Figure 2 Zonally averaged mean observed minus background radiance departures for the AIRS long-wave channels (15 micron) ranked vertically

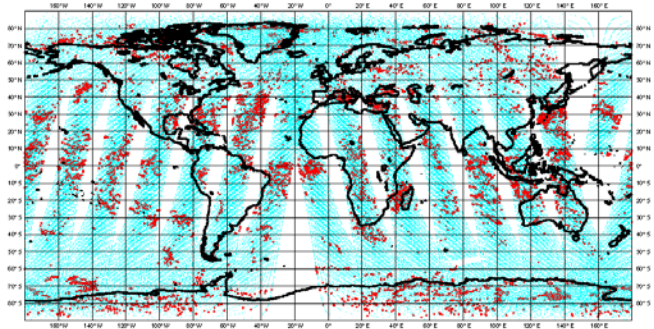
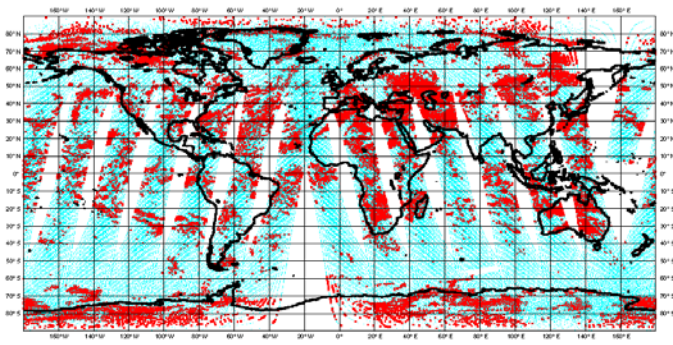
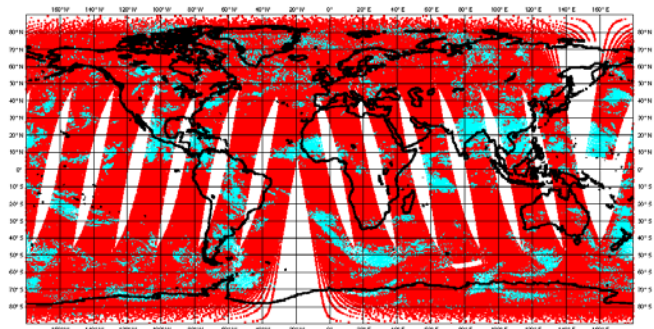


Figure 3 Location of clear data (red symbols) for a lower stratospheric channel at 14.5 microns (top panel), a mid-tropospheric channel at 13.5 microns (centre panel) and a window channel at 11 microns (lower panel)

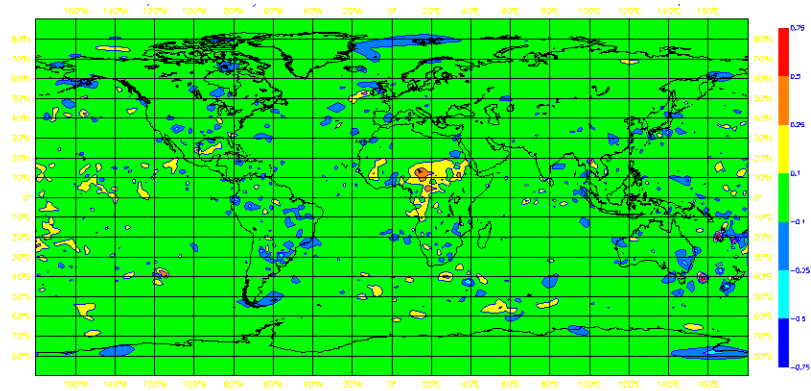


Figure 4 Difference map showing RMS analysis increments of the AIRS system minus those of the CTRL for temperature at 500hPa (averaged over 10 days). Shading starts at 0.1K.

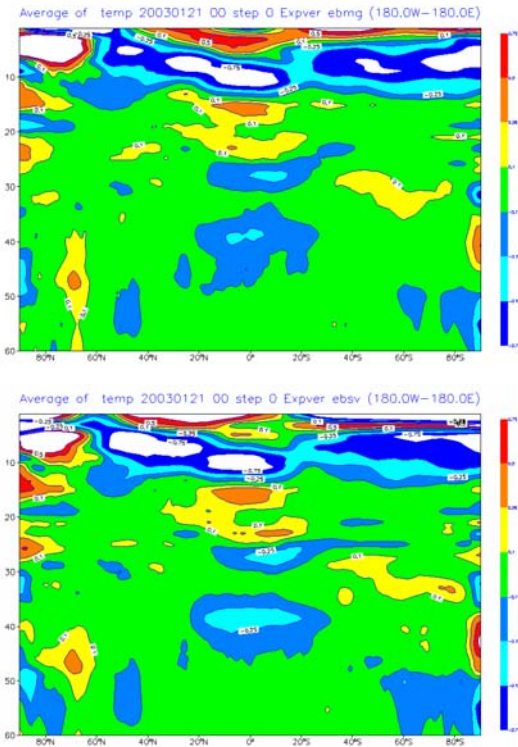


Figure 5 Zonally averaged mean temperature increments (evaluated over 10 days) for the CTRL (top panel) and AIRS (lower panel). Shading starts at 0.1K

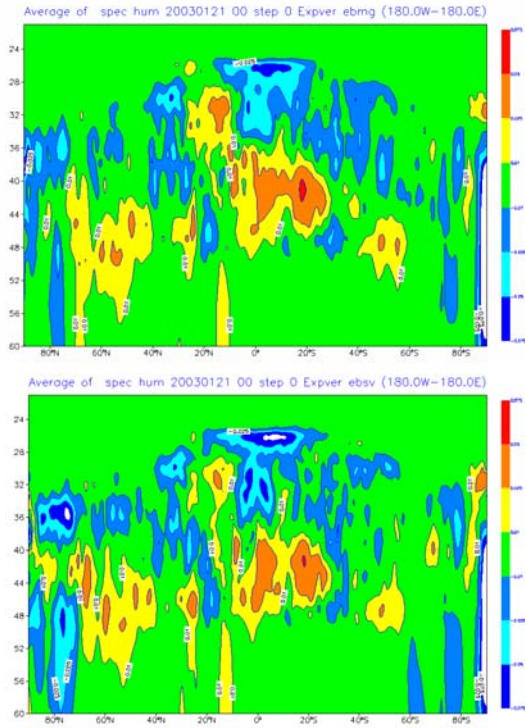


Figure 6 Zonally averaged mean humidity increments (evaluated over 10 days) for the CTRL (top panel) and AIRS (lower panel). Shading starts at 0.01 ($\Delta Q / Q$)

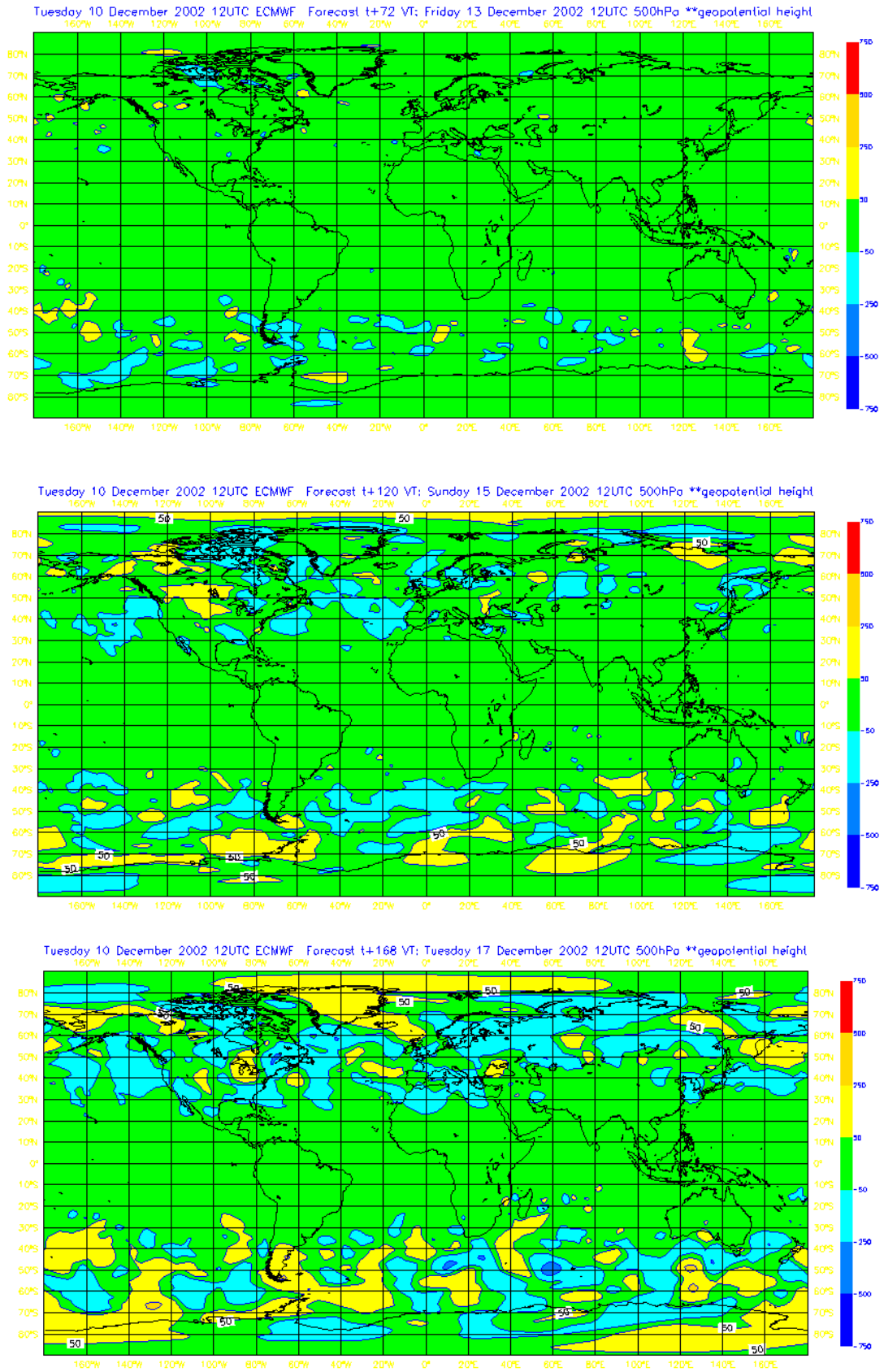


Figure 7 RMS forecast error differences for 500hPa height AIRS minus CTRL