Effects of data selection on the assimilation of AIRS data

Popular Summary

The Atmospheric InfraRed Sounder (AIRS) is one of five instruments flying on NASA's Aqua satellite that was launched in 2002. It measures heat emitted by the Earth's atmosphere, surface, and clouds. It makes these measurements in 2378 separate channels. It maps the Earth twice daily at approximately 1:30 AM and PM. Each day, approximately 35 Gigabytes of AIRS data are processed and archived.

A primary goal of the AIRS mission is to improve weather forecasts. In order for AIRS data to be useful for this purpose, it must be transmitted from the satellite, processed, and distributed to forecasting centers within three hours of the time the measurements are made. Computer systems that bring together all the available data for weather forecasts simply cannot process all of the AIRS data even if it could be distributed in time. Therefore, only a portion of AIRS data are sent to forecast centers.

There are several ways to reduce the volume of AIRS data. One way is to send only a small subset of the 2378 channels. Another is to thin out the mapped data. This leads to several questions such as 1) How do we pick which channels to distribute and use? 2) How much weight do we give to the different AIRS channels as compared with other instruments as we combine all the information together? 3) What is the best way to thin the data? Here, we compare forecast accuracy using different subsets of AIRS channels, weighting schemes, and thinning methods. Our results surprised us. We expect that using more channels will lead to more improvement in forecast skill. However, we find that adding certain channels actually degrades forecasts. There are many possible explanations for this. We had to run several more experiments to rule the possibilities in or out. Finally, we found a viable explanation. AIRS data provide information about both temperature and humidity throughout much of the Earth's atmosphere. This information is all mixed up in the AIRS channels. Even though there are thousands of channels, it is extremely difficult to sort it all out. We have to rely on statistical information, which has large uncertainties, to try to disentangle the mixture. It's like having coffee and orange juice for breakfast. Both are good drinks by themselves. However, if you mix them together, it does not make for a tasty drink, and it is difficult indeed to get back the original good parts. Another surprising result was that we did not see significant improvement in forecast skill when we used more sophisticated (better) methods of thinning the data.

Finally, we consider the overall improvement to forecasts using our best set of approximately 150 AIRS channels. We find that AIRS yields a small but significant improvement in forecast skill when used in combination with the 15 channel Advanced Microwave Sounding Unit (AMSU) that also flies on the Aqua satellite. This improvement extends useful forecast skill by several hours. Decades ago, when AIRS was first conceived, its designers thought it would give a much larger improvement. What happened? During the time it takes to build and launch a new instrument, instruments such as AMSU developed and were launched on other satellites. These instruments all give small improvements to weather forecasts. For example, AMSU's strength is that it can see through most types of clouds, whereas AIRS cannot. So even though it does not have that many channels, it provides a similar improvement to forecasts. As forecasts incrementally improve with each instrument, it becomes more and more difficult to further increase their skill. So the next time you turn to your 5 day weather forecast with confidence, please remember that it takes a combination of many different instruments (including those from NASA satellites) along with powerful computers to bring it to you.

by Joanna Joiner et al.

Effects of data selection on the assimilation of AIRS data

Popular Summary

The Atmospheric InfraRed Sounder (AIRS) is one of five instruments flying on NASA's Aqua satellite that was launched in 2002. It measures heat emitted by the Earth's atmosphere, surface, and clouds. It makes these measurements in 2378 separate channels. It maps the Earth twice daily at approximately 1:30 AM and PM. Each day, approximately 35 Gigabytes of AIRS data are processed and archived.

A primary goal of the AIRS mission is to improve weather forecasts. In order for AIRS data to be useful for this purpose, it must be transmitted from the satellite, processed, and distributed to forecasting centers within three hours of the time the measurements are made. Computer systems that bring together all the available data for weather forecasts simply cannot process all of the AIRS data even if it could be distributed in time. Therefore, only a portion of AIRS data are sent to forecast centers.

There are several ways to reduce the volume of AIRS data. One way is to send only a small subset of the 2378 channels. Another is to thin out the mapped data. This leads to several questions such as 1) How do we pick which channels to distribute and use? 2) How much weight do we give to the different AIRS channels as compared with other instruments as we combine all the information together? 3) What is the best way to thin the data? Here, we compare forecast accuracy using different subsets of AIRS channels, weighting schemes, and thinning methods. Our results surprised us. We expect that using more channels will lead to more improvement in forecast skill. However, we find that adding certain channels actually degrades forecasts. There are many possible explanations for this. We had to run several more experiments to rule the possibilities in or out. Finally, we found a viable explanation. AIRS data provide information about both temperature and humidity throughout much of the Earth's atmosphere. This information is all mixed up in the AIRS channels. Even though there are thousands of channels, it is extremely difficult to sort it all out. We have to rely on statistical information, which has large uncertainties, to try to disentangle the mixture. It's like having coffee and orange juice for breakfast. Both are good drinks by themselves. However, if you mix them together, it does not make for a tasty drink, and it is difficult indeed to get back the original good parts. Another surprising result was that we did not see significant improvement in forecast skill when we used more sophisticated (better) methods of thinning the data.

Finally, we consider the overall improvement to forecasts using our best set of approximately 150 AIRS channels. We find that AIRS yields a small but significant improvement in forecast skill when used in combination with the 15 channel Advanced Microwave Sounding Unit (AMSU) that also flies on the Aqua satellite. This improvement extends useful forecast skill by several hours. Decades ago, when AIRS was first conceived, its designers thought it would give a much larger improvement. What happened? During the time it takes to build and launch a new instrument, instruments such as AMSU developed and were launched on other satellites. These instruments all give small improvements to weather forecasts. For example, AMSU's strength is that it can see through most types of clouds, whereas AIRS cannot. So even though it does not have that many channels, it provides a similar improvement to forecasts. As forecasts incrementally improve with each instrument, it becomes more and more difficult to further increase their skill. So the next time you turn to your 5 day weather forecast with confidence, please remember that it takes a combination of many different instruments (including those from NASA satellites) along with powerful computers to bring it to you.

by Joanna Joiner et al.

Q. J. R. Meteorol. Soc. (2004), **128**, pp. 1–999 doi: 10.1256/qj.

Effects of data selection on the assimilation of AIRS data

By J. JOINER[1]*, E. BRIN[2], R. TREADON[4], J. DERBER[4], P. VAN DELST[4], A. DA SILVA[3], J. LE MARSHALL[4], P. POLI[5,6], R. ATLAS[7] C. CRUZ[8], D. BUNGATO[2]

[1]*NASA Goddard Space Flight Center Laboratory for Atmospheres, USA*
[2]*Science Applications International Corporation, USA*
[3]*NASA Goddard Space Flight Center Global Modeling and Assimilation Office, USA*
[4]*NOAA National Centers for Environmental Prediction, USA*
[5]*Joint Center for Earth Systems Technology, University of Maryland, Baltimore County, USA*
[6]*Now at Centre National de Recherches Météorologiques, Météo France, Toulouse, France*
[7]*NOAA Atlantic Oceanographic and Meteorological Laboratory, USA*
[8]*NGITI, USA*

(Received 1 January 2006; revised 31 January 2006)

SUMMARY

The Atmospheric InfraRed Sounder (AIRS), flying aboard NASA's Earth Observing System (EOS) Aqua satellite with the Advanced Microwave Sounding Unit-A (AMSU-A), has been providing data for use in numerical weather prediction (NWP) and data assimilation systems (DAS) for over three years. The full AIRS data set is currently not transmitted in near-real-time (NRT) to the NWP centers. Instead, data sets with reduced spatial and spectral information are produced and made available in NRT. In this paper, we evaluate the use of different channel selections and error specifications. We achieved significant positive impact from the Aqua AIRS/AMSU-A combination in both hemispheres during our experimental time period of January 2003. The best results were obtained using a set of 156 channels that did not include any in the 6.7μm water vapor band. The latter have a large influence on both temperature and humidity analyses. If observation and background errors are not properly specified, the partitioning of temperature and humidity information from these channels will not be correct, and this can lead to a degradation in forecast skill. We found that changing the specified channel errors had a significant effect on the amount of data that entered into the analysis as a result of quality control thresholds that are related to the errors. However, changing the channel errors within a relatively small window did not significantly impact forecast skill with the 155 channel set. We also examined the effects of different types of spatial data reduction on assimilated data sets and NWP forecast skill. Whether we picked the center or the warmest AIRS pixel in a 3×3 array affected the amount of data ingested by the analysis but had a negligible impact on the forecast skill.

KEYWORDS: forecast numerical weather climate radiances satellite

1. INTRODUCTION

The Atmospheric Infra-Red Sounder (AIRS) Sounding Unit-A (AMSU-A) (Aumann *et al.* 2003) is the first of several advanced high-spectral-resolution nadir-viewing passive infrared sounders to be used for climate applications and operational numerical weather prediction (NWP). AIRS is a grating spectrometer that has been flying on the National Aeronautics and Space Administration's (NASA) Earth Observing System (EOS) polar-orbiting Aqua platform since May 2002 along with the Advanced Microwave Sounding Unit - A (AMSU-A) and several other instruments. Over the next few years, several kilochannel interferometers will fly in Low Earth Orbit. These include the Infrared Atmospheric Sounding Interferometer (IASI) on the EUMETSAT MetOp platform and the Cross-Track Infrared Sounder (CrIS) on the National Polar-orbiting Operational Environmental Satellite System (NPOESS) series of satellites as well as the NASA/National Oceanic and Atmospheric Administration (NOAA)/(US) Department of Defense (DoD) NPOESS Preparatory Project (NPP).

In order to facilitate near-real-time (NRT) transmission of the voluminous AIRS data, the complete AIRS data set must be reduced. There are several possible methods of data reduction. These include channel and/or pixel subsetting

* Corresponding author: NASA Goddard Space Flight Center, Code 613.3, Greenbelt, MD, 20771, USA.

TABLE 1. AIRS 281 channel subset used here in terms of the original 2378 channels. Channels not used for any experiments are shown in italics.

1	6	7	10	11	15	16	17	20	21	22	24	27
28	30	36	39	40	42	51	52	54	55	56	59	62
63	68	69	71	72	*73*	*74*	*75*	*76*	*77*	*78*	*79*	*80*
82	*83*	*84*	86	92	93	98	99	101	104	105	108	110
111	113	116	117	123	124	128	129	138	139	144	145	150
151	156	157	159	162	165	168	169	170	172	173	174	175
177	179	180	182	185	186	190	192	198	201	204	207	210
215	216	221	226	227	232	252	253	256	257	261	262	267
272	295	299	300	305	310	321	325	333	338	355	362	375
453	475	484	497	528	587	672	787	791	843	870	914	950
1003	1012	1019	1024	1030	1038	1048	1069	1079	1082	1083	1088	1090
1092	1095	1104	1111	1115	1116	1119	1120	1123	1130	1138	1142	1178
1199	1206	1221	1237	1252	1260	1263	1266	1285	1301	1304	1329	1371
1382	1415	1424	1449	1455	1466	1477	1500	1519	1538	1545	1565	1574
1583	1593	1614	1627	1636	1644	1652	1669	1674	1681	1694	1708	1717
1723	1740	1748	1751	1756	1763	1766	1771	1777	1780	1783	1794	1800
1803	1806	1812	1826	1843	1852	1865	1866	1868	1869	1872	1873	1876
1881	1882	*1883*	1911	1917	1918	1924	1928	*1937*	*1941*	*2099*	*2100*	*2101*
2103	*2104*	*2106*	*2107*	*2108*	*2109*	*2110*	*2111*	*2112*	*2113*	*2114*	*2115*	*2116*
2117	*2118*	*2119*	*2120*	*2121*	*2122*	*2123*	*2128*	*2134*	*2141*	*2145*	*2149*	*2153*
2164	*2189*	*2197*	*2209*	*2226*	*2234*	*2280*	*2318*	*2321*	*2325*	*2328*	*2333*	*2339*
2348	*2353*	*2355*	*2357*	*2363*	*2370*	*2371*	*2377*					

and methods such as principle component analysis that represent only the most important modes of the spectral information content. Before launch, the NOAA National Environmental Satellite Data and Information Service (NESDIS) set up a special processing system to provide several different data sets to the NWP and data assimilation community (Goldberg *et al.* 2003).

Shortly after launch, the NWP and data assimilation centers began to receive subsetted AIRS NRT data. Positive impact on NWP skill has been demonstrated (*e.g.,* Le Marshall *et al.* 2005, McNally *et al.* 2006) and AIRS data assimilation is now operational at a few centers.

This paper examines the impact on AIRS assimilation of using several channel selections, assigning different channel weights, and selecting a particular type of spatial subsetting. Section 2 describes the AIRS instrument and data preprocessing in detail. The fvSSI data assimilation system (DAS) and experimental setup are discussed in sections 3 and 4, respectively. Results from several assimilation experiments are presented in section 5. A discussion focused on the effect of one particular set of channels is given in section 6. Conclusions and suggestions for further research are provided in section 7.

2. AIRS INSTRUMENT AND DATA PRE-PROCESSING

AIRS is a cooled array grating spectrometer with 2378 channels covering the spectral range 3.7-15.4μm (650-2400 cm^{-1}) with a resolving power ($\nu/\Delta\nu$) of 1200. The instrument scans cross track over a swath width of 1650 km. The footprint diameter is approximately 13.5 km at nadir from the nominal orbital height of 705 km. which corresponds to an aperture of 1.1° . The reduced data sets used here contain a 281 channel subset of the 2378 available AIRS channels. These channels are listed in table 1.

Figure 1 shows normalized weighting functions for AIRS channels centered at 704.4, 1524.4, and 1045.3 cm^{-1}. These channels are affected primarily by CO_2,

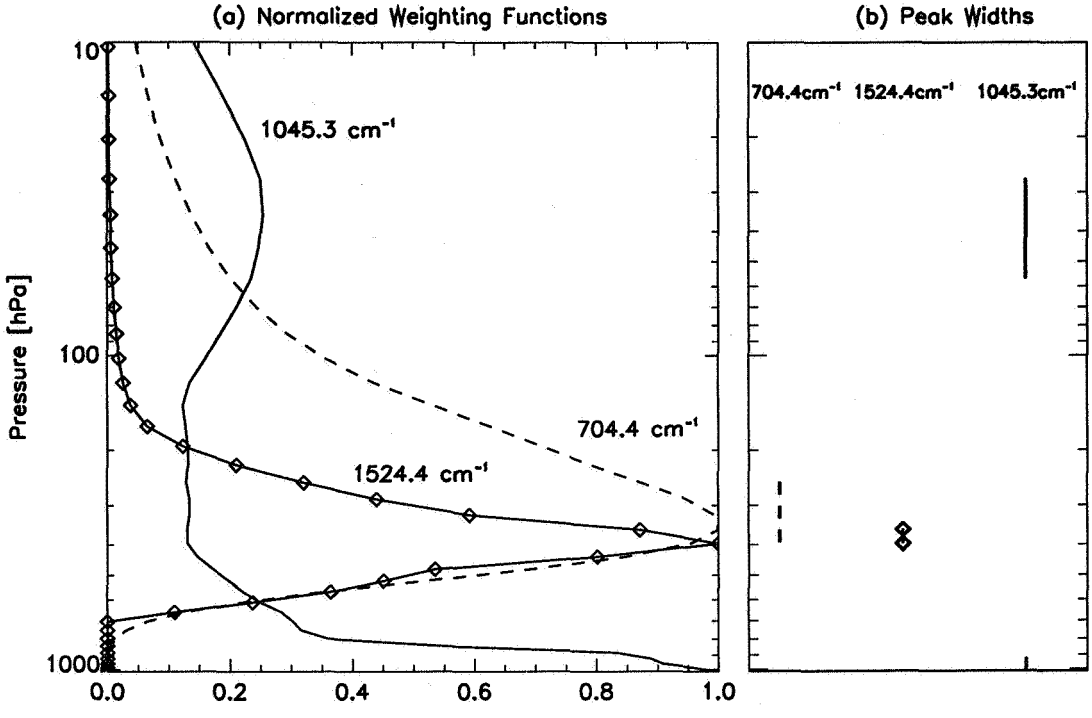

Figure 1. a) Normalized weighting functions for AIRS channels centered at 704.4, 1524.4, and 1045.3 cm^{-1}. b) Weighting function peaks and widths (see text for more explanation) for the same channels.

H_2O, and O_3 absorption, respectively. The weighting functions were computed using a midlatitude profile. The 1045.3 cm^{-1} channel has two peaks, one due to ozone absorption in the stratosphere and one due primarily to water vapor absorption in the lower troposphere. The weighting function widths shown in figure 1b are defined as the range for which the function is at or greater than 90% of its peak value for a given peak.

There are nine AIRS pixels within a collocated AMSU-A footprint. This combination is known as a "golfball." In the cross-track direction, there are 30 golfballs (30 AMSU-A and 90 AIRS pixels). The AIRS and AMSU-A radiance data set used here retains half of the available golfballs. Most NWP centers currently receive only the center AIRS pixel of a golfball. However, we receive in near-real-time and use here a data set that includes all nine AIRS pixels within a golfball. This allows us to evaluate different methods of spatial data reduction.

3. The fvSSI Data Assimilation System

We use a data assimilation system (DAS) that will be referred to as the fvSSI. The fvSSI uses the general circulation model (GCM) of Lin *et al.* (2004), called the fvGCM. The fvGCM consists of the finite volume (fv) dynamical core of Lin (1997) with NCAR CCM3 physics (Kiehl *et al.* 1996). The model top is 0.01hPa. The gridpoint fvGCM is run at a horizontal resolution of 1° latitude × 1.25° longitude in both forecast and data assimilation modes.

Figure 2 shows a simulated AIRS brightness temperature spectrum with the 281 channel subset highlighted. It can be seen in the CO_2 absorption band (*e.g.*, between ~720-740 cm^{-1}) and in the H_2O absorption band (*e.g.*, between ~1300-1400 cm^{-1}), most of the channels in the 281 subset are those in the wings of absorption lines. These lines typically have sharper weighting functions than those near line centers. However, in order to get the highest peaking channels for water vapor, the 281 channel subset also includes channels on the line centers of the strongest lines (*e.g.*, near 1550 cm^{-1}). Also shown for reference are the weighting function peaks and widths.

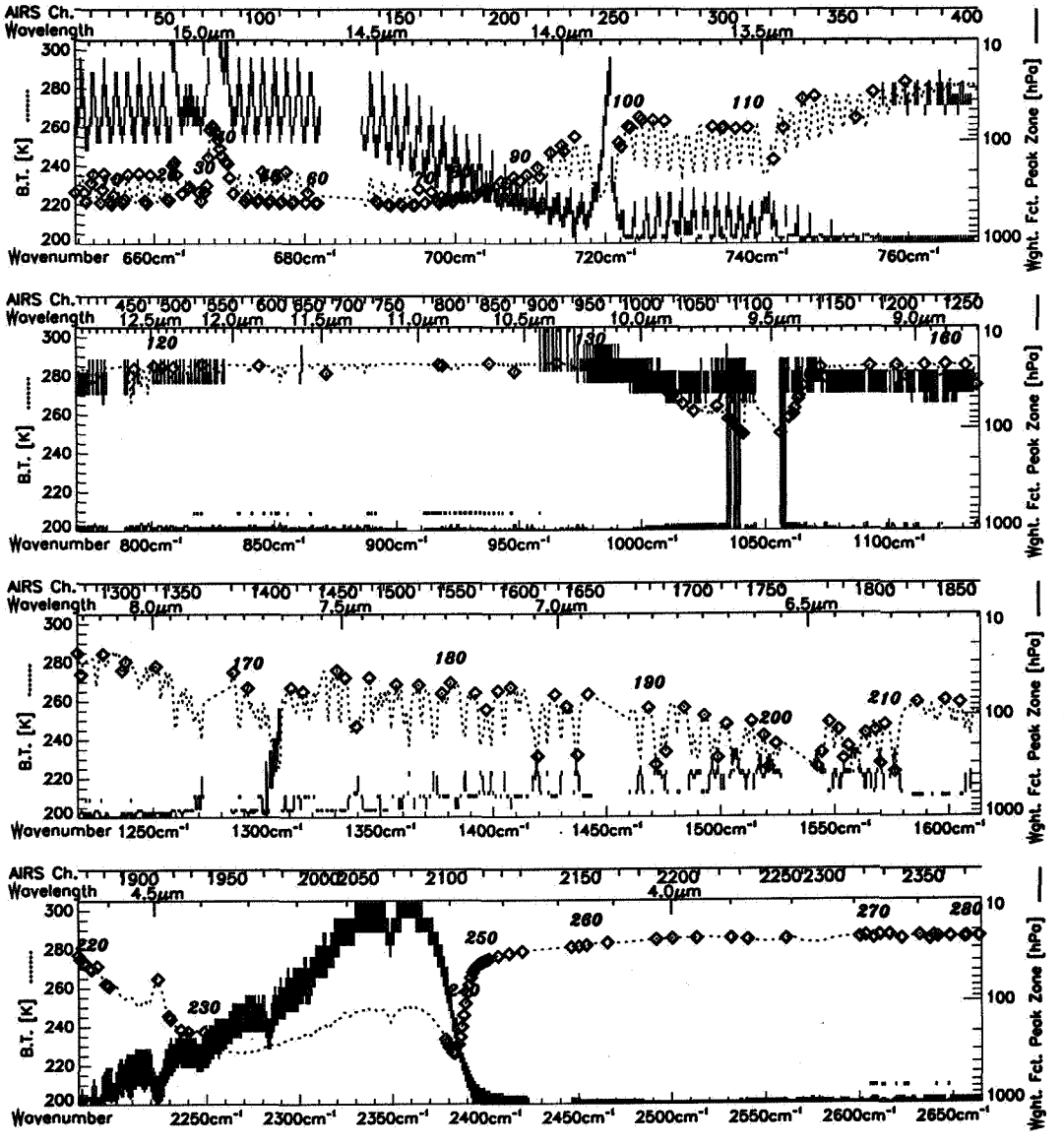

Figure 2. Simulated AIRS brightness temperature spectrum (dotted line). ◇: channels from the 281 subset with the channel number listed directly above. The topmost horizontal scale is the channel number in terms of the full 2378 channel set (labeled AIRS ch.); The lefthand scale is the brightness temperature (B.T.) in Kelvins; The vertical lines give the peaks and extents of the channel weighting functions in terms of pressure on the righthand scale as in figure 1.

The analysis system is the 3D variational (3DVAR) Spectral Statistical Interpolation (SSI) analysis developed at the National Center for Environmental Prediction (NCEP) (Parrish and Derber 1992; Derber and Wu 1998). In order to conduct a relatively large number of experiments, we use a lower horizontal resolution (T62L64: spectral triangular truncation of 62 (~200km) and 64 vertical levels) than the operational NCEP analysis system. For our experiments, we froze a version of the SSI that is now older than the one currently operational at NCEP. It therefore does not include more recent updates that have further improved forecast skills (see http://www.emc.ncep.noaa.gov/gmb/STATS/html/model_changes.html for more details). Our version evolved from one that was operational circa 2004 but includes some updates such as a more recent version of the radiative transfer code and updated observational errors for satellite radiances.

The input observations consist of most of the data that were operationally assimilated at NCEP during the time-period evaluated here: January 2003. These include conventional data such as radiosonde temperatures, humidities, and winds. Cloud-track, water vapor, and ocean surface winds from several satellites are used. In addition, fvSSI assimilates ozone retrievals from the NOAA 16 Solar Backscatter Ultraviolet (SBUV) radiometer. There is no precipitation assimilation in the fvSSI as there is currently in NCEP's operational system.

Our system also assimilates brightness temperatures from the NOAA 14 High Resolution Infrared Sounder 2 (HIRS-2) and MSU, NOAA 15 AMSU-A and AMSU-B, NOAA 16 HIRS-3, AMSU-A, and AMSU-B, the Geostationary Operational Environmental Satellite (GOES) 8 and 10 sounders, and the NASA EOS Aqua AMSU-A and AIRS as described in McNally *et al.* (2000) and Derber and Wu (1998). Note that AIRS data were not assimilated operationally in January 2003. Brightness temperatures are computed within the DAS using the Community Radiative Transfer Model (CRTM), formerly known as OPTRAN (Kleespies *et al.* 2004). The bias correction scheme includes scan-position-dependent corrections for each channel of every instrument. We set the initial bias correction coefficients for all channels to zero at the beginning of the experiment. We then specify a time constant parameter that controls how quickly the system adjusts to the computed bias corrections. We use a value of two days in order to obtain a relatively rapid response to the observations.

One relatively new feature of the SSI analysis system is the cloud detection scheme used for radiance data. This scheme finds a cloud fraction and cloud top pressure that best agrees with the radiance data from an individual sounding. Then, the analysis uses only those channels that peak above the cloud top. This allows the DAS to use potentially many unaffected AIRS channels over a cloudy pixel.

4. Experimental setups

Table 2 gives a summary of the names and experimental parameters for all experiments. We provide an explanation of these parameters in the following subsections.

(*a*) *EOS Aqua channel selection and assigned channel errors*

Each AIRS channel is assigned a constant brightness temperature error for a given experiment. We conducted experiments with two different sets of channels

TABLE 2. Experimental setup (see text for more details). Alternate experiment names are given in parentheses. The last four columns list the number of channels falling within the given spectral ranges.

Experiment name	AMSU used	spatial subset	AIRS chan. errors	H_2O cycled	650-920 cm^{-1}	920-1080 cm^{-1}	1080-1610 cm^{-1}	2180-2242 cm^{-1}	650-2242 cm^{-1}
Control (AMSU)	Yes	-	-	Yes	-	-	-	-	-
No H_2O (Small) (Warmest FOV)	Yes	warmest	small	Yes	115	29	0	12	156
No H_2O, no O_3	Yes	warmest	small	Yes	115	0	0	12	127
H_2O (Small)	Yes	warmest	small	Yes	115	29	59	12	215
No H_2O Large†	Yes	warmest	large	Yes	115	29	0	12	156
H_2O* Large	Yes	warmest	large	Yes	85	4	49	14	152
No AIRS/AMSU	No	-	-	Yes	-	-	-	-	-
AIRS	No	warmest	small	Yes	115	29	0	12	156
Center FOV	Yes	center	small	Yes	115	29	0	12	156
No cycle H_2O	Yes	warmest	small	No	115	29	59	12	215

†Only 19 forecasts available

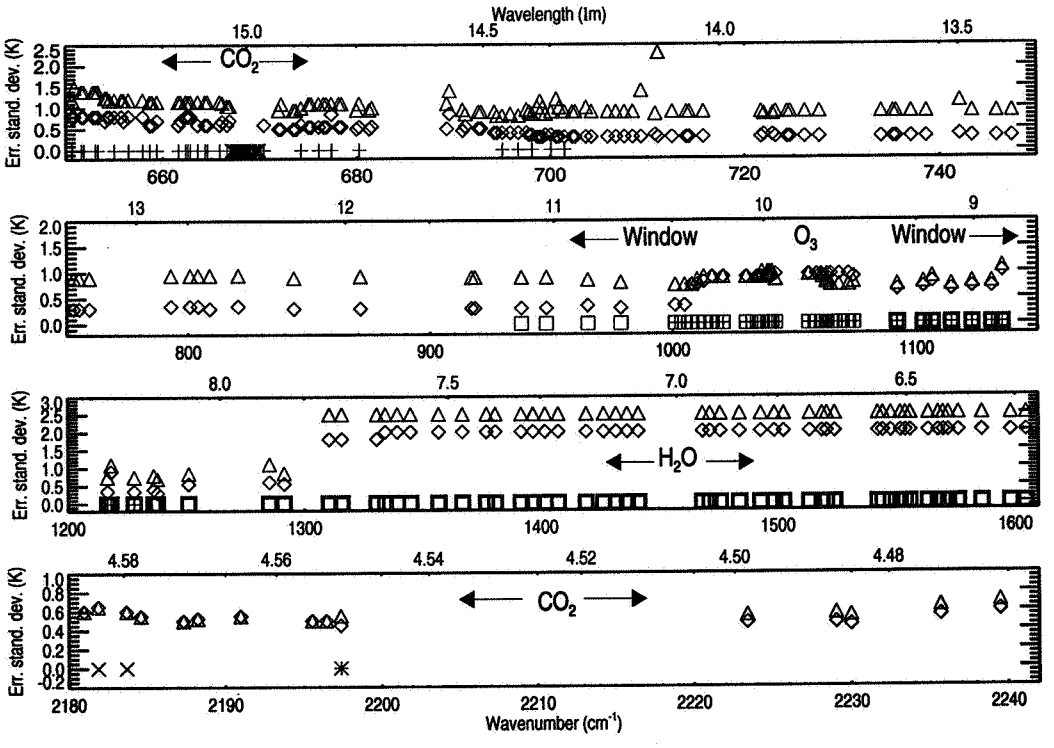

Figure 3. Assigned brightness temperature errors for AIRS channels. △: Large error set; ◇: Small error set; Channels that are not used in given experiments are shown as having zero error. ×: not used in Small set; +: not used in H_2O* large; thick squares: not used in No H_2O (Small and Large); thin square: not used in No H_2O, no O_3 (in addition to thick squares). Major absorbers and band characteristics are indicated.

errors shown in figure 3. The first set, referred to as Small, assigns errors of approximately 0.3K to channels between about 700 and 920 cm^{-1}. Slightly larger errors are assigned to channels with $\nu <\sim 700$ cm^{-1} and those in the 4μm CO_2 band. We give channels in the 6.7μm water vapor band 2K errors. The second set, called Large, has errors of 0.7K in the 690-1300cm^{-1} range with 2.5K errors in the 6.7μm water vapor band. In the 4μm CO_2 band, the Large channel errors are very similar to those of Small.

Using these two sets of channel errors, we then conduct experiments with different selections of channels. The channels that are eliminated in different experiments are indicated in figure 3 by symbols plotted at 0K. For example, in the experiment called No H_2O Small, all channels between 1080 and 1610 cm^{-1} are eliminated (shown as thick squares). In addition to those eliminated in No H_2O Small, the experiment called No H_2O, no O_3 also discards channels between 920 and 1080 cm^{-1} as indicated by the thin squares. One of the experiments that used the Large errors (H_2O* Large) discards channels between 990 and 1240 cm^{-1} as well as a few channels near 700 $cm^{-1}$and several of the channels that peak near the tropopause between 650 and 680 cm^{-1}.

Channels in the strongest part of the 15μm CO_2 band near 667 cm^{-1} are not used in any of the experiments. These channels have weighting functions that peak in the upper stratosphere with tails in the mesosphere. In this work, we also discard all channels shortward of 4.44μm. These include either somewhat redundant window channels and 4.3μm CO_2 band channels that may have somewhat larger uncertainties in spectroscopy due to line mixing. In addition, some of the 4.3μm CO_2 channels are affected by reflected solar radiation and/or non-local thermal equilibrium (non-LTE) during the day.

For the EOS-Aqua AMSU-A, channel 7 is not used due to excessive noise. In addition, as with all other AMSU-A instruments, channel 14 is not used. This channel is sensitive to temperature in the upper stratosphere and mesosphere. All other channels are used. However, channels 1-3 and 15 are assigned errors of 5K or more and therefore receive little weight.

The specified channel errors affect the SSI analysis in two important ways. Firstly, they determine how much weight a particular channel will receive. Secondly, a strict quality control check is performed whereby an observation is not used if the absolute value of the observed minus background (henceforth referred to as O-B) brightness temperature is greater than either three times the specified error standard deviation or 4.5K. Note that *e.g.,* if we want to give less weight to a particular channel by increasing the error, we must consider that more data from the channel will enter the analysis owing to the larger quality control threshold. This could potentially allow more cloud-contaminated data into the analysis.

(*b*) *Spatial subsetting*

We perform assimilation experiments with two different methods of spatial subsetting. Specifically, we choose one of the nine AIRS fields-of-view (FOV) within a golfball by selecting either the center FOV or the warmest FOV (as defined by the brightness temperature in the 917.1cm^{-1} channel). The advantage of the center FOV selection is that it potentially simplifies the across-track bias correction, because the same scan positions will be used for all scan lines. The warmest FOV selection can potentially find more holes in the clouds. However, the warmest FOV procedure does not always ensure that the most channel data

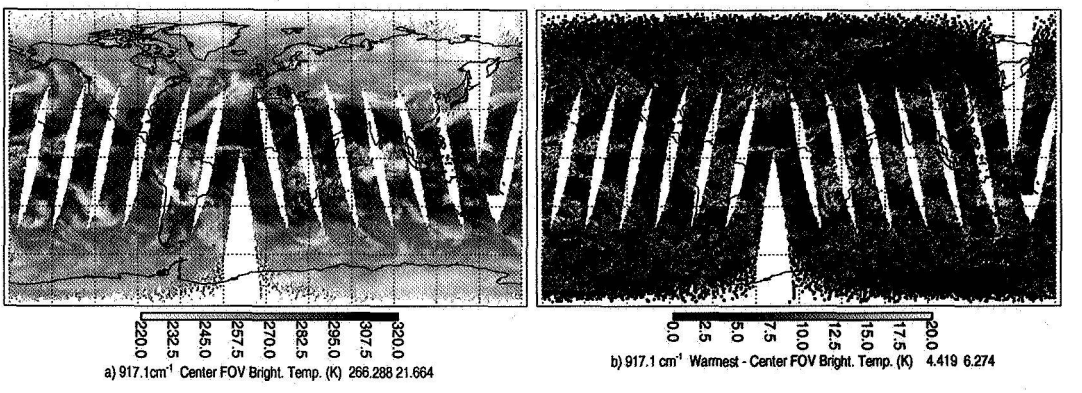

Figure 4. a) Observed brightness temperatures for AIRS channel centered at 917.1cm^{-1} for 20 December 2002, 03Z-15Z, center FOV subsetting. b) Difference between warmest and center FOV subsetted data sets. The numbers at the bottom right are the global means and standard deviations of the brightness temperature, respectively.

for a given golfball will be used or that the least cloudy pixel will be selected. For example, a high cloud with either a low emissivity or low cloud fraction may produce a warmer brightness temperature in a window channel than a lower cloud with a higher emissivity/cloud fraction. A temperature inversion may produce cloudy brightness temperature observations that are warmer than surrounding clear ones.

Figure 4a shows brightness temperature observations for the AIRS window channel centered at 917.1 cm^{-1} on 20 December 2002 using the center FOV subsetting. Brightness temperature differences for the same channel between the warmest and center FOV spatial subsettings are shown in figure 4b. The largest differences occur in cloudy regions. These are precisely the areas where forecasts may have a large sensitivity to AIRS data (McNally 2002; Fourrie and Rabier, 2004). The mean brightness temperature between the two data sets is significant at more than 4K while the standard deviation is also large (over 6K).

5. Results of AIRS assimilation experiments

(a) *Data coverage and O-B statistics*

Figure 5 shows the percentage of the thinned data that pass the cloud detection and background checks for the warmest FOV (Small and Large errors) and center FOV (Small errors). The statistics were computed for 06Z on 20 December 2002. Both the spatial subsetting and the specified channel errors play significant roles in determining how much data enters the analysis. As expected, the warmest FOV subsetting allows more data to enter the analysis for channels peaking in the lower troposphere that are affected by cloud. Changing the channel errors can have an even larger effect on the amount of data accepted by the analysis owing to the fact that the errors determine the thresholds for the background quality control check.

Figure 6 similarly shows global O-B statistics for the same three experiments. The method of spatial subsetting has little effect on either the O-B means or standard deviations. Even though the center-FOV observations are cooler on average (for the entire AIRS population), the O-B statistics for the two spatial

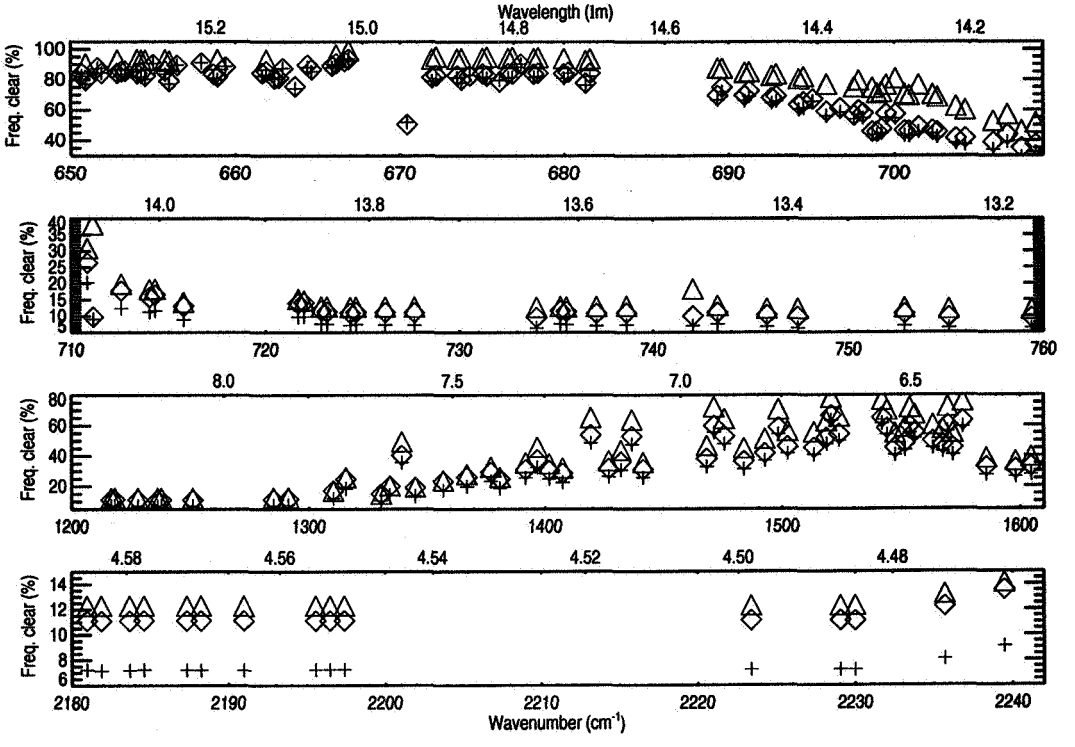

Figure 5. Percentage of thinned AIRS data that passed the analysis cloud detection and quality control checks: △: warmest FOV, No H_2O Large; ◇: warmest FOV, No H_2O Small; +: center FOV, No H_2O Small. Some channels are shown as if they had been used in the analyses, but were not actually used. Other channels have been omitted (*e.g.*, some channels in the 11μm window) as their data counts appear very similar to other channels already shown.

subsets are similar because a smaller population of AIRS pixels is determined to be clear for the center-FOV case. However, the specified channel errors do have a significant effect on the O-B statistics, particularly the standard deviation for channels with $\nu < 1000$ cm^{-1}. This is expected as the channel errors affect the quality control thresholds. Note that there is no difference in standard deviations in the 6.7μm water vapor band. Because the quality control threshold is capped at 4.5K which is less than 3σ for these channels, changing the channel errors from 2 to 2.5K did not impact quality control decisions. For the lowest peaking channels in the 11-12μm window, the bias is larger and more negative when using the Large channel errors. This could be an indication that some cloud-contaminated data enters the system when we specify the Large channel errors.

Figure 7 shows maps of O-B and coverage for 12 hours of observations (same time period as in figure 4). Here, we show data for the 704.4 cm^{-1} CO_2 channel and for the 1524 cm^{-1} water vapor-sensitive channel. These channel weighting functions can peak at similar pressures (~336 hPa). As shown in figure 5, more observations (~ 10%) are accepted in the warmest FOV subsetting (figure 7b) as compared with the center FOV subsetting (figure 7a) mostly at the edges of cloud-covered areas and in partly cloudy regions. An illustration of this is seen over southern Africa.

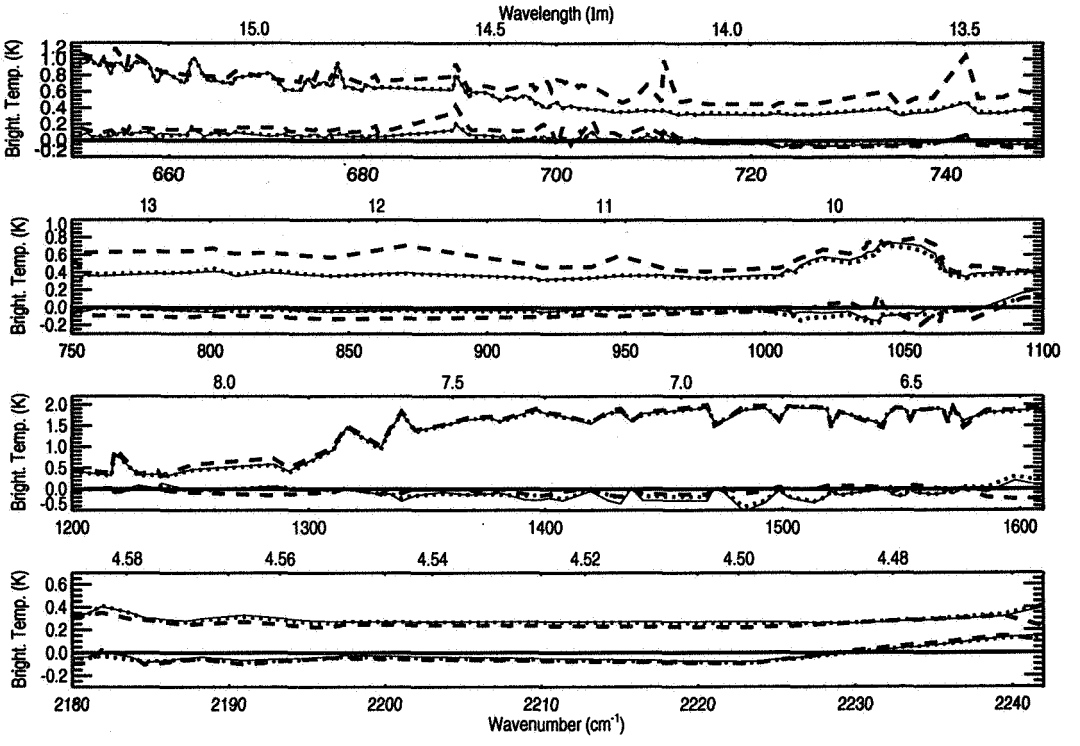

Figure 6. As in figure 5 but showing globally-averaged O-B statistics: Mean (lower set of lines near the thick solid zero line) and standard deviations (upper set of lines). Solid: Warmest FOV (No H_2O Small); Dotted: Center FOV (No H_2O Small); Dashed: H_2O Large (Warmest FOV).

There is a large increase in the number of accepted observations ($\sim 40\%$) when the Large errors are used (figure 7b-c). This increase is especially apparent at high latitudes. In the northern high latitudes, the O-B tends to be positive which is inconsistent with cloud-contamination. However, at southern high latitudes, O-B is more often negative. Figure 4 indicates that these latitudes can be frequently cloud-covered, although this channel may sometimes peak above low cloud. It is difficult to determine from this figure whether cloud-contaminated data are entering the analysis.

Even more observations from the 1524.4 cm^{-1} water vapor channel enter the analysis (figure 7d) as compared with the 704.4 cm^{-1} CO_2 channel whose weighting function peaks at a similar altitude. Reasons for this include the different error specifications that affect quality control decisions and the fact that the 1524.4 cm^{-1} weighting function is more variable and is more narrow (see figure 1) so that it may more frequently peak above low clouds. For example, there is better coverage over low clouds in the middle and high southern latitudes. The O-B for this channel shows smaller-scale structure presumably due to background humidity errors.

(*b*) *Cloud detection*

Here, we compare the analysis cloud-top pressures with those from the EOS Aqua Moderate-Resolution Imaging Spectroradiometer (MODIS). The MODIS

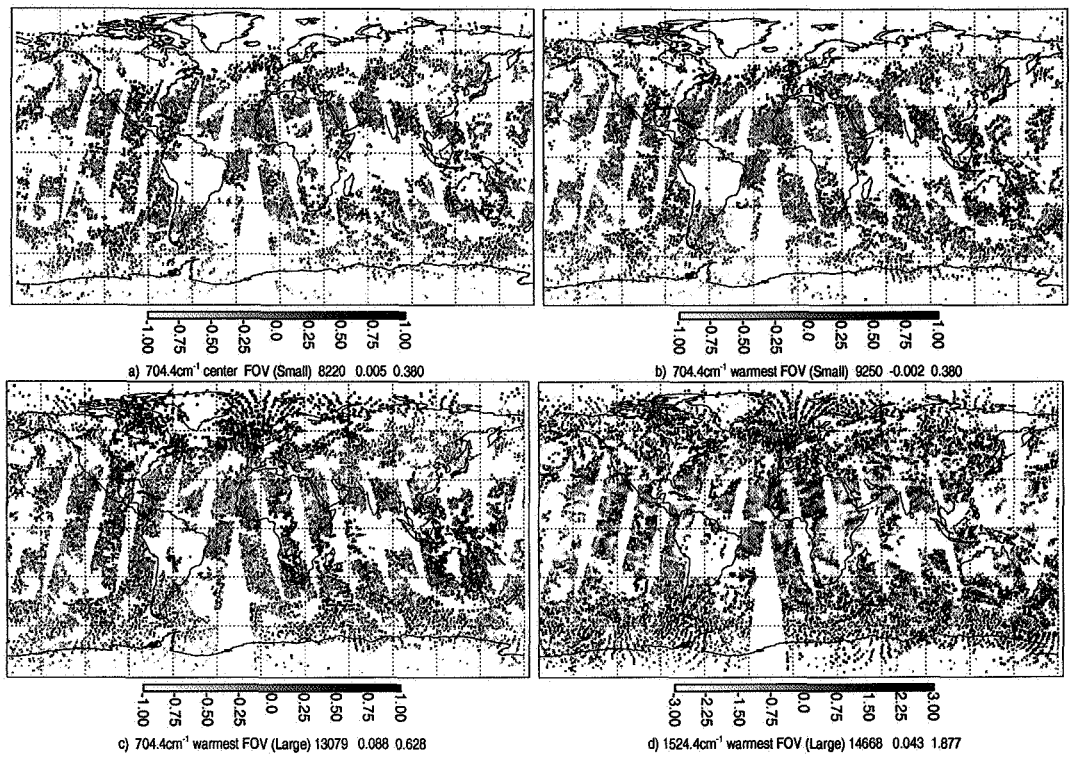

Figure 7. a) O-B for AIRS channel at 704.4cm^{-1} for 20 December 2002, 03Z-15Z, center FOV subsetting (No H_2O Small); b) Same as a) but warmest FOV subsetting; c) Same as b) but for H_2O^* Large; d) Same as c) but AIRS channel at 1524.4 cm^{-1} (note different scale). The numbers at the bottom of each subfigure are the total number of accepted observations, global means, and standard deviations of the O-Bs, respectively.

cloud-top pressures are derived using a CO_2 slicing technique with 1km^2 pixels (Menzel and Strabala, 1997). We use the MODIS cloud-top pressure contained in the level-3 atmosphere product (MOD08) collection 4. The level-3 data are statistics (*e.g.,* mean, minimum, maximum) that are sorted into 1° latitude × 1° longitude cells on an equal-angle global grid. The cloud-top pressures are separated into daytime only, nighttime only, or combined day and night. Here, we use the separate products to match up orbits with AIRS. At latitudes above 60°, we use the combined day-night data set.

Figure 8a-c shows the MODIS minimum, maximum and fvSSI cloud pressures. Multiple cloud decks with holes in the upper layer are present in gridboxes that have a large difference between MODIS cloud-top minima and maxima. AIRS may or may not be able to see through these holes. The fvSSI generally does a good job of differentiating between deep convective clouds in the tropics and lower frontal clouds at higher latitudes. The success of the cloud detection algorithm is less certain over very warm land surfaces such as in southern Africa and Australia. MODIS short-wave (SW) infrared channels indicate high fractions of cirrus cloud (figure 8d) over these locations while the fvSSI sometimes finds either no or low clouds. Joiner *et al.* (2004) explained a potential difficulty in using an O-B-type approach for cloud detection in such areas. A significant background error in the surface skin temperature can offset cloud effects in O-B. This may result

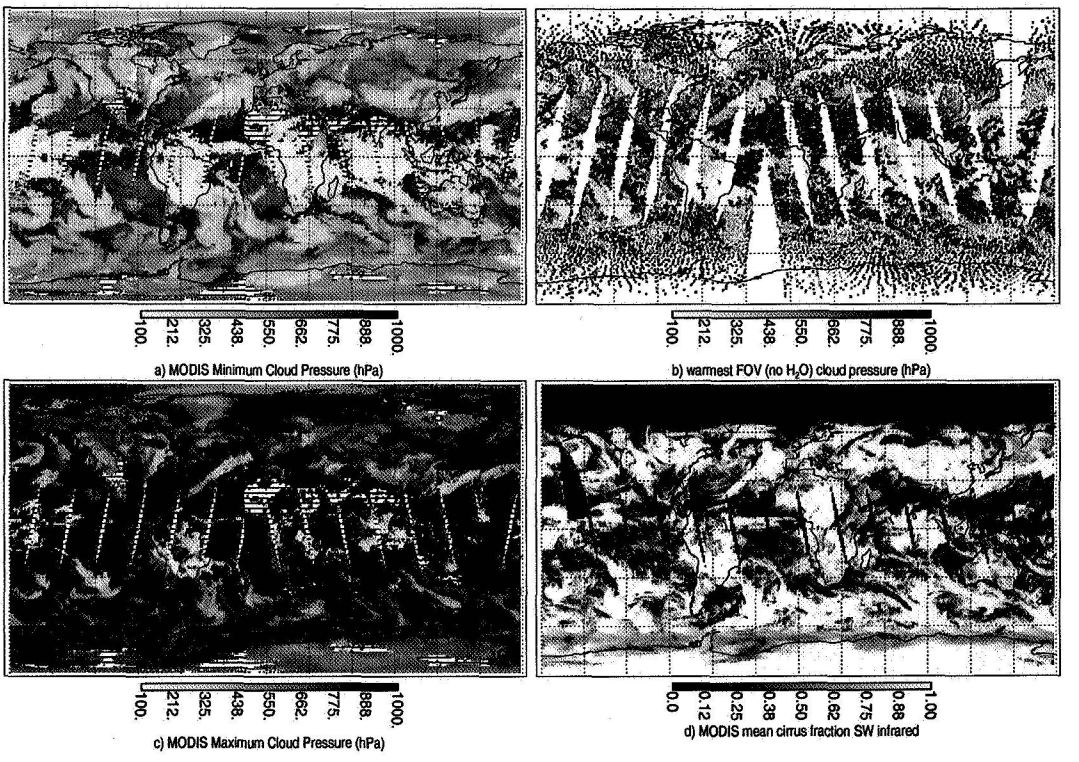

Figure 8. Cloud parameters from MODIS and fvSSI DAS for 20 December 2002: a) MODIS gridbox minimum cloud pressure; b) cloud pressure from fvSSI DAS (03-15Z); c) MODIS gridbox maximum cloud pressure; d) cirrus fraction from MODIS (day orbits only). Hatched areas in a) and c) indicate gridboxes with either no observations or where the cloud fraction was zero. Cloud pressures in b) are shown as the surface pressure where the cloud fraction was set to zero.

in a cloud-contaminated observation passing cloud-detection checks. Additional cloud-detection tests may be needed to improve results under such conditions.

(c) *Forecast skill*

All experiments begin with a two week spin-up period beginning in mid-December 2002. The forecasts are run daily for January 2003. The forecast results for each experiment are verified against the operational NCEP analyses at $2° \times 2.5°$ resolution. This may slightly penalize the results with AIRS data as AIRS data were not assimilated at this time. The forecast anomaly correlation scores include all waves. The extratropical scores are averaged (area weighted) over latitudes from 30° to 86°. A summary of all the scores is given in table 3 along with confidence levels for differences between relevant pairs of experiments. The details are discussed below.

(i) *Channel selection*

Figure 9 shows forecast anomaly correlation scores for four experiments. The Control uses no AIRS data and three others use different AIRS channel selections. All experiments with AIRS data have a positive impact in the northern hemisphere as compared with the Control. No H_2O Small has the largest impact. In the southern hemisphere, No H_2O Small has a neutral impact and the other

TABLE 3. Summary of 500 hPa geopotential height anomaly correlation (AC) scores for the northern and southern hemispheres (NH, SH, respectively) and and the statistical confidence levels for differences from the control (C1) and no H_2O Small (C2). Statistically significant results (to the 90% level) are shown in italics.

Experiment	Cont	no H_2O small	H_2O small	H_2O* large	no H_2O no O_3	no H_2O large	no AIRS AMSU	AIRS	center FOV	no cycle
AC NH	.824	.833	.828	.830	.831	.821†	.821	.829	.833	.824
AC SH	.764	.763	.757	.750	.760	.780†	.754	.749	.757	.760
C1 NH	-	*.985*	.646	*.938*	*.966*	*.999*†	.729	.713	*.993*	.049
C1 SH	-	.057	.737	*.950*	.402	.216†	*.939*	*.931*	.811	.344
C2 NH	*.985*	-	*.984*	.792	*.946*	.049†	*.995*	*.909*	.034	*.994*
C2 SH	.057	-	*.955*	*.904*	.702	.659†	.759	*.995*	.825	.462

†19 forecasts

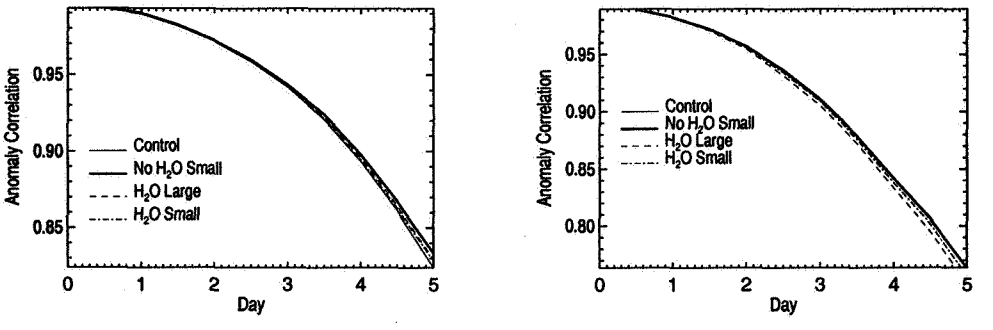

Figure 9. 31 day average of forecast 500 hPa anomaly correlation for northern (left) and southern (right) hemispheres.

experiments with AIRS data have a slightly negative impact. Note that very similar results (in terms of the magnitude of improvement) were obtained by McNally *et al.* 2006) using a higher resolution DAS during an experiment that spanned our time period.

We find that removing 9.7μm channels in addition to the 6.7μm channels had little effect. Using the Large errors compared with the Small similarly had little impact on forecast skill even though this had a significant impact on the amount of data accepted by the analysis. Note that due to loss of the computational platform used for all experiments here, we have a somewhat smaller sample for the No H_2O Large experiment. This sample was compared to the same from the Control and No H_2O Small to generate the confidence levels in table 3.

We see from these experiments that the inclusion of channels in the 6.7μm H_2O band slightly degrades the AIRS impact. This degradation could be the result of one or more effects. Firstly, if the GCM has significant biases in the humidity field, then assimilating good humidity information may actually degrade the assimilation (*e.g.*, Chen *et al.* 1999) by *e.g.*, producing excessive convection and precipitation.

Secondly, channels in the 6.7μm water vapor band have a large sensitivity to both temperature and humidity. These channels can also have very sharp weighting functions and may enter the analysis more often than similarly peaking CO_2 channels over low clouds and as a result of quality control decisions. If the

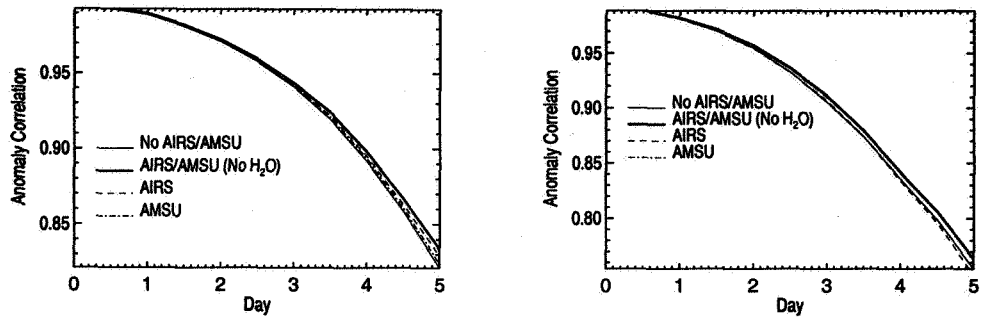

Figure 10. Similar to figure 9 but for AIRS (no H_2O small) and EOS AMSU-A combination experiments.

background errors for temperature and humidity are misspecified, this may result in incorrect partitioning of temperature and humidity increments for a given brightness temperature innovation.

Finally, channels may cause degradation if quality control thresholds are inadequate or if the bias correction scheme is insufficient to account for systematic errors in the forward model or observations themselves. We focus on the first two of these underlying effects of the 6.7μm channels in more detail in section 6.

(ii) *Instrument selection*

Figure 10 examines the impact of the AIRS (No H_2O Small from above) and the EOS Aqua AMSU-A separately and together as compared with an experimental run that used neither AIRS nor EOS-AMSU-A. In the northern hemisphere, AIRS gives a positive impact without the EOS-AMSU-A and a larger impact than the EOS-AMSU-A. In this hemisphere, the use of both instruments together provides a larger positive impact than the sum of the two alone. However, the results are somewhat different in the southern hemisphere. Here, AIRS alone gives a neutral to slightly negative impact while the EOS-AMSU-A gives a positive impact. The addition of AIRS to the EOS-AMSU-A does not yield significant improvement in this hemisphere.

It should be noted that there is some redundancy in the orbits of the EOS Aqua and NOAA-16. Therefore, the impact from the EOS Aqua AIRS/AMSU-A combination may be somewhat less than if it were in a completely independent orbit.

(iii) *Spatial subsetting*

As shown above, the method of spatial subsetting had a significant effect on the amount of data ingested into the analysis. There are indications of improvement using the warmest FOV selection in the southern hemisphere with a confidence level of about 80%. However, this is not generally considered to be a statistically significant result.

6. Discussion

In order to examine the effect of interaction between the moisture analysis and the model physics, we reconfigured our data assimilation system to analyse

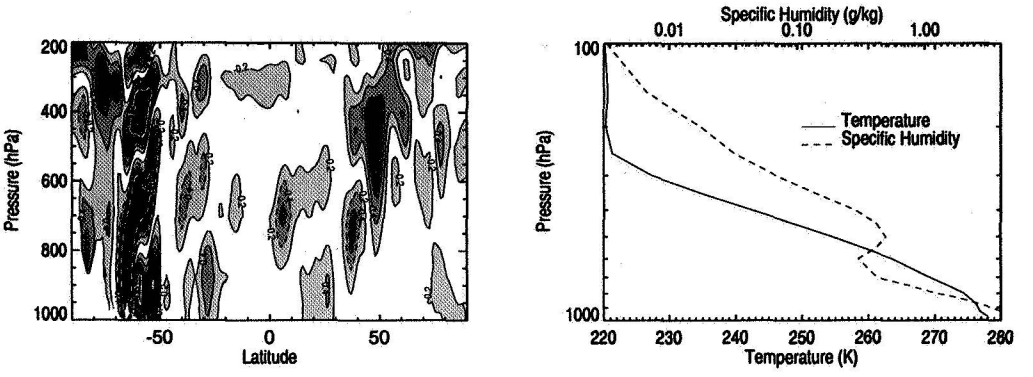

Figure 11. Left: Cross-section through 122.5°W of difference in temperature increments (K) between analyses with and without 6.7μm AIRS channels (only AIRS and EOS-AMSU-A radianaces were assimilated). Contour lines are drawn at 0.2K increments up to 1K with increasingly dark shades (differences > 1K are shown as 1K). Negative contours are shown with dashed lines. Right: Background temperature and specific humidity profiles for 60°S, 122.5°W.

humidity as usual (with all instruments), but then to not feed back the analysed moisture field to the model. This allows humidity-sensitive channels in the 15μm CO_2 band have the potential benefit of a simultaneous moisture analysis with all other instruments. At the same time, the possibility of negative interaction between the humidity analysis and the model physics is eliminated. The model essentially runs with its internally-generated humidity field.

We ran the system in this configuration (called No cycle H_2O) with the same channel set as in H_2O Small. We find that the hemispherically-averaged 5-day forecast skills are very similar to those in the standard configuration with the same AIRS channel set (H_2O Small). Given this result, it is likely that other factors are contributing to the degradation of skill stemming from the use of channels in the 6.7μm H_2O band.

In order to isolate the effect of the 6.7μm water vapor band channels on the temperature analysis, we ran a set of experiments with combinations of only AIRS and EOS AMSU-A channels (*i.e.,* no other data). We examine only the first analyses from these experiments on 17 December 2002 00Z.

Figure 11 shows a cross-section through 122.5°W longitude of the difference in temperature increments between analyses with and without 6.7μm channels. Both experiments use EOS-AMSU-A data and the channel selection and errors of H_2O and No H_2O Small. The largest differences in temperature increments are at middle to high latitudes. Around 60°S, an oscillating vertical pattern is present. There were a number of clear soundings near this gridbox where the analysis ingested all of the AIRS channel data. The background temperature and specific humidity profiles for 60°S, -122.5°W are also shown in figure 11. We next examine the increments at this location in more detail.

Figure 12 shows temperature and specific humidity increments at this location for four different channel configurations. Figure 13 shows differences in increments with respect to the experiment with all AIRS (no AMSU-A) channels. AMSU-A is seen to have a relatively small effect on both the temperature and humidity increments. In contrast, the 6.7μm water vapor channels have a relatively large effect on temperature increments. The overall structure of the temperature increments with all AIRS channels is similar to that produced using only 6.7μm

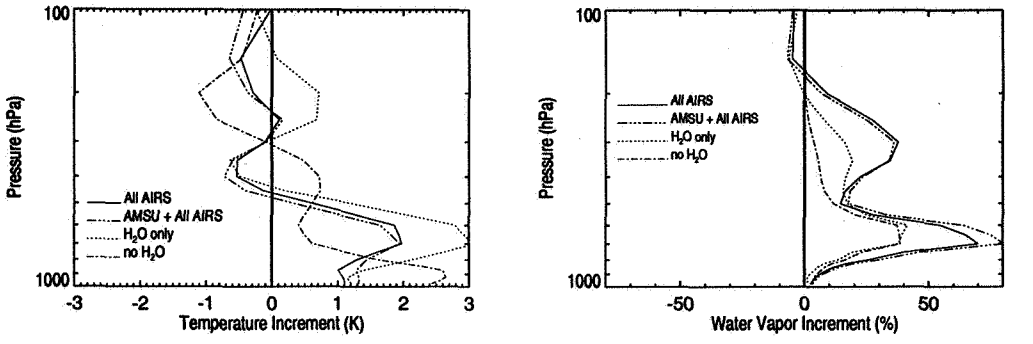

Figure 12. Left: Temperature increments at 60°S and 122.5°W; Right: As in left but specific humidity increments.

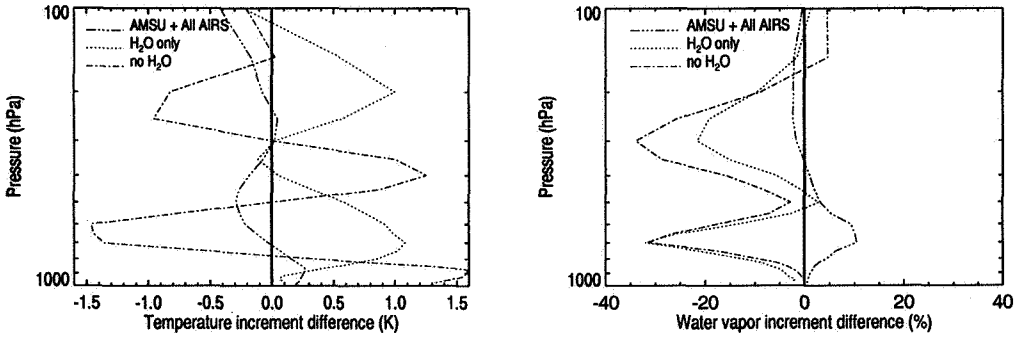

Figure 13. Left: Difference in temperature increments with respect to All AIRS from figure 12; Right: As in left but for specific humidity.

channels and contrasts significantly with the increments generated without those channels. The magnitude of the differences in temperature increments ($\pm$1.5K) produced by the 6.7μm channels is quite significant in this case.

Figure 14 shows spectra of brightness temperatures computed from the model and O-B for a sounding determined to be cloud free that influenced the increments shown above (at -60.55°S, -121.61°W). The generally negative values of O-B near 15μm produce cooling near and above the tropopause ($\sim$200 hPa) as shown in the temperature increments of figure 12. Positive O-B values for $\nu > \sim 750\text{cm}^{-1}$ produce warming in the lower troposphere.

O-B values in the 6.7μm band are positive(negative) for the channels with strong(weak) H_2O absorption. The negative(positive) O-B values should produce an increase(decrease) in water vapor and/or a cooling(warming) in the lower(upper) troposphere. The primary effect of the 6.7μm band O-Bs on the increments is an increase in humidity peaking near 700 hPa. When 6.7μm channels are removed from the analysis, other channels with lower humidity sensitivity (*e.g.*, in the 11μm window) produce a much smaller increase in humidity around 700 hPa, and the warming at this altitude is less.

The vertical oscillations in the temperature increments resulting from the 6.7μm channels are produced by a complex relationship between background

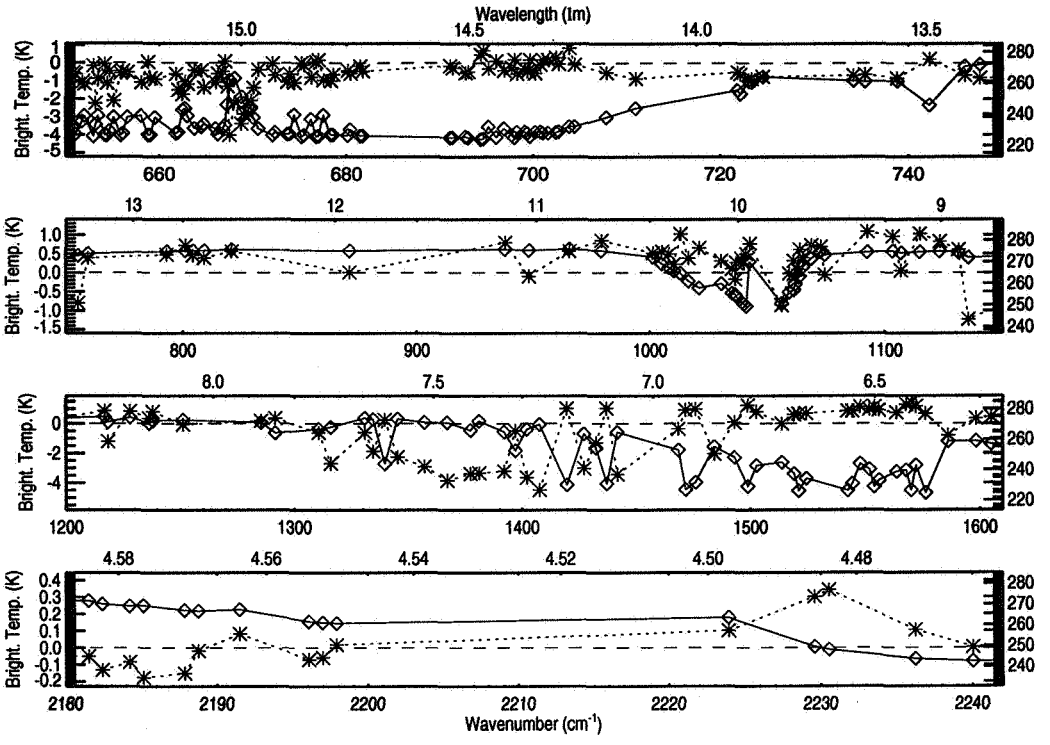

Figure 14. Brightness temperatures computed from the model (solid line with diamonds, right scale) and O-B (stars with dotted line, left scale). The zero line is shown for reference as a dashed line.

errors and the O-B spectra that is not readily apparent from the O-B spectra alone. We do not know exactly where the truth lies in this case. Information content studies show that 6.7μm channels do provide useful information about temperature as well as humidity and complement information from the CO_2 bands (*e.g.,* Rabier *et al.* 2002). However, based on our forecast skill scores, this information is likely being at least partially incorrectly partitioned in our DAS.

To make use of the information content in the 6.7μm H_2O band, background errors as well as observation errors have to be correctly specified. Both the background error variances and vertical correlations play a role in determining how the increments will be spread over the different state variables and throughout the vertical. Because background errors vary both spatially and temporally, they are very difficult to accurately estimate.

With a more simple observable that is a function of only type of state variable (*e.g.,* radiosonde temperature measurements), a good analysis can be achieved if the ratio of the background error to the observation error is correctly specified. A scalar weight can be defined for a given observation based on this ratio or vice-versa. In other words, it is not necessary to have the correct absolute magnitudes of the both the observation and background errors. However, with satellite radiances that are sensitive to both temperature and humidity, the situation has an added level of complexity. The concept of a scalar weight cannot be used. The ratio of the temperature and humidity errors projected onto brightness

temperature space and its relationship to observation errors becomes important. This ratio is much more difficult to estimate.

This problem, of course, applies to all sounders that have sensitivity to both temperature and humidity (as well as other state variables). These include HIRS, MSU, and AMSU-A, all of which have <20 sounding channels. AIRS has a large number of relatively low noise channels with many in the 6.7μm H_2O band that have highly peaked weighting functions and frequently peak above low clouds. Therefore, it is likely that the problem of correctly partitioning the increments is amplified with AIRS as compared with the other sounders.

We obtained our best result by simply omitting 6.7μm channels from the analysis. However, there may be a more optimal set of channel errors and quality control thresholds that will provide a better use of these channels. These parameters may be determined experimentally, but the computational cost to do this is high. We did not have the resources to run additional experiments with other sets of errors.

7. Summary and Future Work

We have achieved a significant positive impact in both hemispheres by using the AIRS and AMSU-A combination of instruments from the EOS Aqua satellite. The best results were obtained using a set of 156 channels that did not include any from the 6.7μm water vapor band. We find that the assigned channel errors and the different methods of spatial subsetting for AIRS both play a significant role in determining the amount of data that is accepted into our fvSSI DAS. However, we also find that ingesting more AIRS radiances in the DAS does not always translate into improved forecasts. For example, the warmest FOV spatial subsetting yielded improvement in terms of data coverage in partially cloudy conditions. However, somewhat to our surprise, although there are indications of a slight improvement with the warmest FOV subsetting in the southern hemisphere, we did not achieve a statistically robust increase in forecast skill with this data selection.

We find that channels in the 6.7μm water vapor band can have a significant impact on temperature analyses. Although these channels in AIRS-type instruments have been shown in simulations to provide useful information for temperature sounding, the simulations also assume that background errors are known, unbiased, and Gaussian. In reality, these assumptions, in particular knowledge of the errors, are likely flawed. Without a very good estimate of the background errors, and especially their variation in both time and space, it is unlikely that current assimilation systems can fully exploit information from this band. In addition, use of these channels can produce a negative result as we have seen here.

We caution that the results obtained here may or may not translate to other data assimilation systems. Furthermore, our experiments are conducted in one season using a relatively low horizontal resolution as compared with the operational NCEP DAS. It will be interesting to see whether similar results are achieved using different data assimilation systems.

In our experiments, the assigned channel errors affect both the weights those channels receive in the analysis and quality control decisions. In the future, we plan to explore decoupling the quality control thresholds from the assigned errors. More sophisticated schemes for spatial subsetting and cloud detection are

also being investigated. Finally, we plan to conduct additional experiments that evaluate the use of cloud-cleared radiances.

Acknowledgements

The authors thank the AIRS science team, particularly M. Chahine and H. H. Aumann for their continued support and enthusiasm, M. Goldberg and W. Wolf for providing timely EOS Aqua AIRS and AMSU-A data, and R. Todling for assistance with the DAS. This work was supported by National Aeronautics and Space Administration (NASA) through the Joint Center for Satellite Data Assimilation (JCSDA) and the Modeling and Assimilation Program (MAP).

References

Aumann, H. H., Chahine, M. T., Gautier, C., Goldberg, M. D., Kalnay, E., McMillin, L. M., Revercomb, H., Rosenkrantz, P. W., Smith, W. L., Staelin, D. H., Strow, L. L., and Susskind, J. 2003 AIRS/AMSU/HSB on the Aqua Mission: Design, science objectives, data products, and processing systems *IEEE Trans. Geosci. Rem. Sens.*, **41**, 253–264

Chen, M., Joiner, J., and Rood, R. B. 1999 Assimilating TOVS Humidity into the GEOS-2 Data Assimilation System. *J. Clim.*, **12**, 2983–2995

Derber, J., and Wu, W.-S. 1998 The use of cloud-cleared radiances in the NCEP SSI analysis system. *Mon. Wea. Rev*, **126**, 2287–2299

Fourrie N., and Rabier, F. 2004 Cloud characteristics and channel selection for IASI radiances in meteorologically sensitive areas. *Quart. J. Roy. Meteor. Soc.*, **130**, 1839-1856

Goldberg, M. D., Qu, Y., McMillin, L. M., Wolf, W., Zhou, L., and Divakarla, M. 2003 AIRS Near-real-time products and algorithms in support of operational numerical weather prediction. *IEEE Trans. Geosci. Rem. Sens.*, **41**, 379–389

Joiner, J. and Rokke, L. 2000 Variational cloud clearing with TOVS data. *Quart. J. Roy. Meteor. Soc.*, **126**, 725–748

Joiner, J., Poli, P., Frank, D., and Liu, H. H. 2004 Detection of cloud-affected AIRS channels using an adjacent-pixel approach. *Quart. J. Roy. Meteor. Soc.*, **130**, 1469–1487

Kiehl, J. T., Hack, J. J., Bonan, G. B., Boville, B. A., Briegleb, B. P., Williamson, D. L., and Rasch, P. J. 1996 Description of the NCAR Community Climate Model (CCM3). *NCAR Technical Note*, NCAR/TN-420+STR, Boulder, CO, 152pp.

King, M. D., Menzel, W. P., Kaufman, Y. J., Tanré, D., Gao, B. C., Platnick, S., Ackerman, S. A., Remer, L. A., Pincus, R., and Hubanks, P. A. 2003 Cloud and aerosol properties, precipitable water, and profiles of temperature and water vapor from MODIS. *IEEE Trans. Geosci. Rem. Sens.*, **41**, 442–458

Kleespies, T.J., van Delst, P., McMillin, L.M., and Derber, J. 2004 Atmospheric Transmittance of an Absorbing Gas. 6. OPTRAN status report and introduction to the NESDIS/NCEP community radiative transfer model. *Appl. Opt.*, **43**, 3103-3109

Lin, S.-J., Atlas, R.M., and Yeh, K.S. 2004 Global weather prediction and high-end computer NASA*Comp. Sci. Eng.*, **6**, 29–35

Lin, S.-J. 1997 A finite-volume integration method for computing pressure gradient forces in general vertical coordinates. *Quart. J. Roy. Meteor. Soc.*, **123**, 1749–1762

Le Marshall, J., Jung, J., Derber, J., Treadon, R., Lord, S.J., Goldberg, M., Wolf, W., Liu, H.C., Joiner, J., Woollen, J., Todling, R., and Gelaro, R. 2005 Impact of Atmospheric Infrared Sounder observations on weather forecasts. *Eos*, **86**, 109–116.

McNally, A. P., Derber, J.C., Wu, W.S., and Katz, B.B. 2000 The use of TOVS level-1B radiances in the NCEP SSI analysis system. Quart. J. Roy. Meteor. Soc., **126**, 689-724

McNally, A. P. 2002 A note on the occurrence of cloud in meteorologically sensitive areas and the implications for advanced infrared sounders *Quart. J. Roy. Meteor. Soc.*, **128**, 2551–2556

McNally, A. P., Watts, P. D., Smith, J. A., Engelen, R., Kelly, G. A., Thépaut, J. N., and Matricardi, M. 2006 The assimilation of AIRS radiance data at ECMWF. *Quart. J. Roy. Meteor. Soc.*, in press

Menzel, P., and Strabala, K. 1997 Cloud top properties and cloud phase algorithm theoretical basis document, NASA document ATBD-MOD-04, Greenbelt, MD, USA, http://modis-atmos.gsfc.nasa.gov

Parrish D. F., and Derber J. C. 1992 The National-Meteorological-Centers Spectral Statistical-Interpolation analysis system. *Month. Weath. Rev.*, **120**, 1747-1763

Rabier, F., Fourrié, N., Chafai, D., and Prunet, P. 2002 Channel selection methods for Infrared Atmospheric Sounding Interferometer radiances. *Quart. J. Roy. Meteor. Soc.*, **128**, 1011-1027.