

Improved methodology for surface and atmospheric soundings, error estimates, and Quality Control procedures: The AIRS Science Team Version-6 retrieval algorithm

Joel Susskind,^a John Blaisdell,^{a,b} Lena Iredell^{a,b}

^aNASA Goddard Space Flight Center, Earth Sciences Division-Atmospheres, Greenbelt, MD, USA, 20771

^bSAIC/NASA, Code 610, 8800 Greenbelt Rd., Greenbelt, MD, USA, 20771

Abstract. The AIRS Science Team Version-6 AIRS/AMSU retrieval algorithm is now operational at the Goddard DISC. AIRS Version-6 level-2 products are generated near real-time at the Goddard DISC and all level-2 and level-3 products are available starting from September 2002. This paper describes some of the significant improvements in retrieval methodology contained in the Version-6 retrieval algorithm compared to that previously used in Version-5. In particular, the AIRS Science Team made major improvements with regard to the algorithms used to 1) derive surface skin temperature and surface spectral emissivity; 2) generate the initial state used to start the cloud clearing and retrieval procedures; and 3) derive error estimates and use them for Quality Control. Significant improvements have also been made in the generation of cloud parameters. In addition to the basic AIRS/AMSU mode, Version-6 also operates in an AIRS Only (AO) mode which produces results almost as good as those of the full AIRS/AMSU mode. This paper also demonstrates the improvements of some AIRS Version-6 and Version-6 AO products compared to those obtained using Version-5.

Keywords: Remote sensing, infrared, clouds, satellites, meteorology

Address all correspondence to: Joel Susskind, NASA Goddard Space Flight Center, Earth Sciences Division-Atmospheres, 8800 Greenbelt Rd., Greenbelt, MD, USA, 20771; Tel: +1 301-286-7210; Fax: +1 301-286-2323; E-mail: Joel.Susskind-1@nasa.gov

1 Introduction

The AIRS Science Team Version-6 retrieval algorithm, now operational at the Goddard DISC, contains many significant improvements compared to the previously operational AIRS Science Team Version-5 retrieval algorithm. Hundreds of scientific papers have been published showing the benefits of using AIRS Version-5 products. A partial list of these publications can be found at <http://airs/jpl.nasa.gov/documents/publications/>.

The basic cloud clearing and retrieval methodologies used in the AIRS Science Team Version-6 retrieval algorithm, including the meaning and derivation of Jacobians, the channel noise covariance matrix, and the use of constraints including the background term, are essentially identical to those of the AIRS Science Team Version-3 algorithm¹, which was developed and tested using simulated AIRS/AMSU observations. Unlike most other retrieval methodologies,

there is no explicit weight given to either an a-priori state or the initial guess. Susskind et al.² described the AIRS Science Team Version-4 retrieval algorithm used by the Goddard DAAC to analyze AIRS/AMSU observations from September 2002 (when the AIRS instrument became stable) through September 2007, two months after AIRS Version-5 processing began. The AIRS Science Team AIRS/AMSU Version-4 retrieval and cloud clearing algorithms included new terms to account for systematic and random errors made in the computation of expected channel radiances for a given geophysical state using the Version-4 AIRS and AMSU Radiative Transfer Algorithm. Version-4 also introduced a Quality Control (QC) concept that generated different QC flags for a given profile as a function of height, and also had separate QC flags related to surface skin temperature. The AIRS Science Team Version-5 retrieval algorithm³ contained many further improvements. The most important improvement in Version-5 retrieval methodology was made in the set of channels used to retrieve the atmospheric temperature profile. In addition, Version-5 developed methodology to generate profile-by-profile, level-by-level, error estimates of temperature profile and to use them for level-by-level QC flags for the atmospheric temperature profile. The AIRS Version-6 AIRS/AMSU retrieval algorithm contains many further improvements in retrieval methodology beyond what was done in Version-5. Foremost among these is a major improvement in the Version-6 retrieval methodology used to determine surface skin temperature and surface spectral emissivity from AIRS observations. There have also been significant improvements to the QC methodology used for different geophysical parameters, the methodology used to generate first guesses for atmospheric and surface parameters, and the methodology used to determine cloud parameters and compute OLR from the AIRS/AMSU observations. Finally, Version-6 also has an alternate “AIRS Only” (AO) processing capability which utilizes only AIRS observations and produces

results which are only slightly degraded from those obtained utilizing both AIRS and AMSU observations. The Version-6 AO processing mode is an important backup to Version-6 because noise performance on some channels of AMSU-A is continuing to degrade, and at some point, use of Version-6 including AMSU-A observations may become impractical.

2 Overview of the Retrieval Methodologies Used in Both Version-5 and Version-6

Fundamental to all versions of the AIRS Science Team retrieval system is the generation of clear column radiances I_{clear} for each AIRS channel i , which are derived products representing the radiance channel i would have seen if the entire 3x3 AIRS Field of Regard (FOR) on which a retrieval is performed were cloud free. I_{clear} is determined for each channel as a linear combination of the observed radiances of that channel in each of the nine AIRS Fields of View (FOV's) contained within the AIRS FOR, with coefficients that are channel independent¹. The retrieved geophysical state X is subsequently determined which, when substituted into the AIRS Radiative Transfer Algorithm, generates an ensemble of computed radiances I_{comp} which are consistent with I_{obs} for those channels i used in the determination of X . Cloud-clearing theory^{4,5} says that to achieve the best retrieval results under more stressing cloud conditions, longwave channels sensitive to cloud contamination should be used only in the determination of the coefficients used in the generation of clear column radiances for all channels, and not be used for sounding purposes. In Version-5³, tropospheric sounding 15 μm CO₂ observations were used only in the derivation of the cloud clearing coefficients, and temperature profiles were derived using I_{obs} in the 4.3 μm CO₂ band as well as in some stratospheric sounding 15 μm CO₂ channels that do not see clouds. This new approach allowed for the retrieval of accurate QC'd values of Q and $T(p)$ under more stressing cloud conditions than was achievable in Version-4. Version-5 also

contained a new methodology to provide accurate case-by-case level-by-level error estimates for retrieved geophysical parameters as well as for channel-by-channel clear column radiances. Thresholds of these error estimates were used in a new approach for the generation of QC flags in Version-5.

The AIRS Version-6 retrieval algorithm has further significant advances over Version-5. The basic theoretical approach used in Version-6 to analyze AIRS/AMSU data is very similar to what was done in Version-5 with one major exception. As in Version-5, the coefficients used for generation of clear column radiances for all channels are determined in Version-6 using observed radiances only in longwave 15 μm and 11 μm channels. Following cloud clearing theory^{4,5}, Version-5 retrieved tropospheric temperatures using only in the AIRS shortwave 4.2 μm CO₂ channels. Version-5 did not follow this principle with regard to the surface parameter retrieval step, in which in both the longwave 8 - 12 μm window region and in the shortwave 4.0 μm – 3.7 μm window region were used together to simultaneously determine surface skin temperature, surface spectral emissivity, and surface bi-directional reflectance of solar radiation. Version-6 uses only window observations in the shortwave window region 4.0 μm – 3.76 μm determine surface skin temperature along with shortwave surface spectral emissivity and shortwave surface bi-directional reflectance. Longwave surface spectral emissivity is retrieved in Version-6 in a subsequent step using values of only in the longwave window region. Another significant improvement found in Version-6 is the use of an initial guess generated using Neural-Net methodology^{6,7} in place of the previously used two regression approach³. These two modifications have resulted in significant improvement in the ability to obtain both accurate temperature profiles and surface skin temperatures under more stressing partial cloud cover conditions.

2.1 Steps in the Version-5 and Version-6 Retrieval Algorithms

Retrievals of all geophysical parameters are physically based and represent states determined for case c that best match a set of clear column radiances for the subset of AIRS channels i used in the retrieval process. Retrievals of geophysical parameters are performed sequentially, that is, only a subset of the geophysical parameters within the state is modified from that of the incoming state in a given step. A GCM forecast is not used in the retrieval procedure, except for use of the forecasted surface pressure p_{surf} as the lower pressure boundary when computing radiances expected for a given geophysical state. In the case of AIRS Only retrievals, a GCM forecast is also used in the specification of surface class over potentially frozen ocean.

In Version-5, the steps in the physical retrieval process were as follows: A start-up procedure, involving use of a cloudy regression followed by a clear regression, was used to generate the initial state. Initial clear column radiances were generated for all channels i using the initial state and the cloud-clearing coefficients which were generated using observed radiances in an ensemble of cloud clearing channels. The state was also used as the initial guess to the physical retrieval process in which AIRS/AMSU observations were used to retrieve: a) surface skin temperature T_s , surface spectral emissivity and surface bi-directional reflectance of solar radiation; b) atmospheric temperature profile $T(p)$; c) atmospheric moisture profile $q(p)$; d) atmospheric ozone profile $O_3(p)$; e) atmospheric CO profile $CO(p)$; f) atmospheric CH₄ profile $CH_4(p)$; and g) cloud properties and OLR. These steps were done sequentially, solving only for the variables to be determined in each retrieval step while using previously determined variables as fixed with an appropriate uncertainty attached to them which

was accounted for in the channel noise covariance matrix used in that step. The objective in each step (a-f) was to find solutions which best match \mathbf{y} for the subset of channels selected for use in that step, bearing in mind the channel noise covariance matrix. Steps a-f were ordered so as to allow for selection of channels in each step which are primarily sensitive to variables to be determined in that step or determined in a previous step, and are relatively insensitive to other parameters. Separation of the problem in this manner allowed for the problem in each step to be made as linear as possible. Step g was performed using a selected set of the observed radiances \mathbf{y}_g and was performed after the surface and atmospheric conditions have been determined.

In Version-6, there are slight modifications to the sequence of steps used in Version-5 because there are two new steps performed in the retrieval sequence. In Version-5, step a) used channels in both the longwave and shortwave window regions and simultaneously solved for surface skin temperature T_s , shortwave surface spectral emissivity ϵ_{sw} , effective shortwave surface spectral bi-directional reflectance ρ_{sw} , and longwave surface spectral emissivity ϵ_{lw} . In Version-6, only shortwave window channels are used in this retrieval step to simultaneously determine T_s , ϵ_{sw} and ρ_{sw} . The longwave surface spectral emissivity ϵ_{lw} is now solved for in a subsequent step using only channels in the longwave window spectral region. This new step is performed after the humidity profile retrieval step because longwave window radiances can be very sensitive to the amount of atmospheric water vapor. In addition, Version-6 contains a new physical retrieval step performed before the surface temperature retrieval step in which T_s is updated from its initial guess value. This additional step is performed only during the day. The steps used in the Version-6 AO (AIRS Only) algorithm are otherwise identical, but no AMSU-A observations are used in the physical retrieval

process or in the generation of the initial state $\epsilon_{s,0}$, which as stated previously, is determined in Version-6 via a Neural-Net methodology^{6,7} rather than by regression as was done in Version-5.

2.2 Improved Version-6 Surface Parameter Retrieval Methodology

In addition to the separation of the retrieval of surface shortwave spectral emissivity and surface longwave spectral emissivity into two separate steps, Version-6 has also improved other details in the retrieval of surface skin parameters. Version-6 uses an improved multiplicative form of the equation to modify the retrieved surface spectral emissivity $\epsilon_{s,k}$ from its initial guess $\epsilon_{s,0}$ rather than the additive form used in Version-5^{1,3}. The same equation is used when solving for $\epsilon_{s,k}$ in the shortwave emissivity retrieval step and also when solving for $\epsilon_{s,k}$ in the longwave surface emissivity step. In Version-6, we treat the variable to be modified as $\epsilon_{s,k}$ and write

$$(1)$$

where there are k_{max} unknowns $\epsilon_{s,k}$ to be solved for and A_k are triangular functions of frequency as in Version-5. Equation 1 is written in this multiplicative form so that $\epsilon_{s,k}$ if all coefficients are equal to zero. In the shortwave surface parameter retrieval step, in which $\epsilon_{s,k}$ is retrieved simultaneously with T_s and ρ , k_{max} is set equal to four, while in the longwave surface emissivity step, k_{max} is set equal to six. A corresponding multiplicative form is also used in Version-6 to modify $\epsilon_{s,k}$ during the day in the shortwave surface parameter retrieval step

$$(2)$$

Therefore, during the day, nine coefficients, one for ΔT_s where ΔT_s is the difference of the retrieved value of T_s from its initial guess $T_{s,0}$, and four values each of A_k and B_k , are solved for in the shortwave surface parameter retrieval step, and five parameters are solved for at night. The initial guess for surface spectral emissivity in both retrieval steps, $\epsilon_{s,0}$ is set equal to the values

found in the AIRS Science Team ocean emissivity model over non-frozen ocean. Over land and frozen ocean, ϵ_{land} is set equal to values interpolated from the $1^\circ \times 1^\circ$ monthly mean MODIS Science Team surface spectral emissivity data set for the year 2008. As in Version-5, ϵ_{land} is initially estimated as being equal to $\epsilon_{\text{land}}^{\text{est}}$, but is then modified in a subsequent retrieval step in Version-6 which is performed immediately prior to the shortwave surface parameter retrieval step. In this step, not performed in Version-5, ϵ_{land} is updated in a one parameter physical retrieval step, using the same channels as in the surface parameter retrieval step, according to

$$(3)$$

where C is a constant which scales ϵ_{land} but does not change its shape. Inclusion of this step is done to help account for the attenuation of incoming solar radiation by partial cloud cover along the path from the sun to the AIRS FOR on which the retrieval is being performed. The values of ϵ_{land} shown in Equation 3 are used as the initial guess $\epsilon_{\text{land}}^{\text{est}}$ in Equation 2. Determination of this constant prior to the full surface retrieval step significantly improved the retrieved values of T_s , ϵ_{land} , and ϵ_{water} determined during daytime.

3 Channels and Functions Used in Different Steps of Version-6

Figure 1 shows a typical AIRS cloud free brightness temperature spectrum and includes the channels used in both Version-6 and Version-6 AO for cloud clearing, as well as in each of the different steps of the AIRS physical retrieval algorithm. The Version-6 channels used in these steps are described in the next sections.

3.1 Cloud Clearing and Temperature Profile Retrieval

Following cloud-clearing theory^{4,5}, coefficients needed to generate clear column radiances for all channels are determined using observations in select longwave channels whose radiances are sensitive to the presence of clouds. Version-6 uses 57 channels to derive the coefficients which

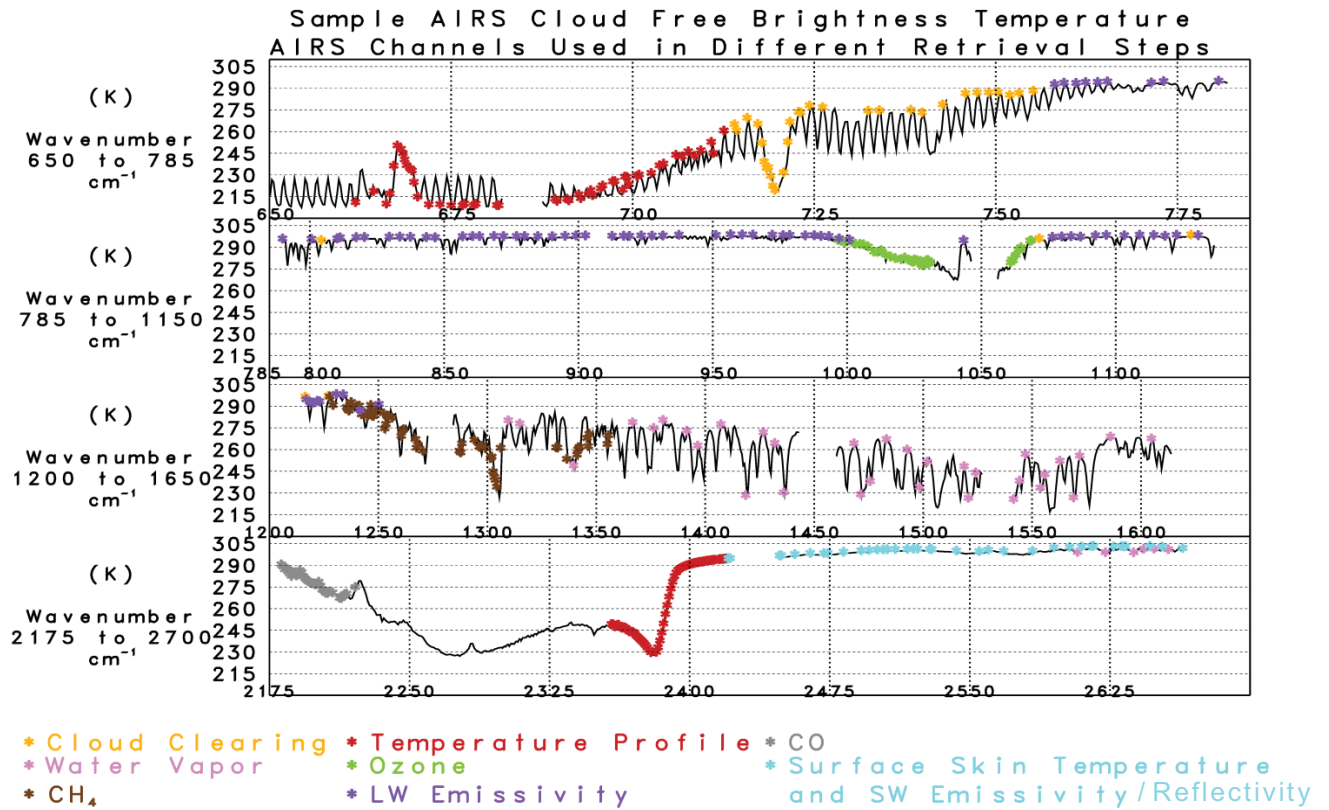


Fig. 1

are used to generate clear column radiances for all channels¹. These channels, which we mark by yellow stars in Fig. 1, range from 701 cm⁻¹ to 1228 cm⁻¹. The cloud clearing channels are the same channels used in a subsequent cloud parameter retrieval step. The temperature profile retrieval step uses 37 channels between 2358 cm⁻¹ and 2395 cm⁻¹ that are sensitive to both stratospheric and tropospheric temperatures, as well as 53 stratospheric sounding channels between 662 cm⁻¹ and 713 cm⁻¹ that are not sensitive to cloud contamination. Longwave channels sensitive to cloud contamination are not used in the temperature profile retrieval step.

We indicate the channels used in the determination of temperature profile by red stars in Fig. 1. Version-6 also includes 24 additional channels in the temperature profile retrieval step between 2396 cm^{-1} and 2418 cm^{-1} , also shown in red, that are used in both the temperature profile step and the surface parameter retrieval step. Version-6 uses AMSU-A channels 3, 6 and 8-14 in the temperature profile retrieval step as well, while Version-6 AO does not use these or any AMSU channels. AMSU-A channel 7 was noisy at launch and was never used in any step of the retrieval process. Version-5 included AMSU-A channels 4 and 5 in the temperature profile retrieval step, but those channels subsequently became noisy and neither is used in Version-6. In addition, Version-5 included 12 AIRS channels between 2198 cm^{-1} and 2252 cm^{-1} in the temperature profile retrieval step that are no longer used in Version-6. These channels are sensitive to absorption by N_2O and were found to contribute to the spurious negative mid-tropospheric temperature trend found in Version-5 because increases in N_2O concentration over time are not accounted for in either the Version-5 or Version-6 retrieval algorithms.

3.2 Surface Skin Temperature and Longwave Spectral Emissivity Retrievals

Unlike in Version-5, the surface skin temperature retrieval and longwave spectral emissivity retrieval are done in separate steps in Version-6. The surface skin temperature retrieval step uses 36 channels between 2420 cm^{-1} and 2664 cm^{-1} , which we show by light blue stars in Fig. 1, along with the 24 highest frequency (red stars) channels which are also used in the temperature profile retrieval step. These 60 channels are used to determine T_s simultaneously with four independent pieces of information about surface shortwave spectral emissivity and, during the day, four additional independent pieces of information about shortwave surface bi-directional reflectance as shown in Equations 1 and 2. Surface longwave spectral emissivity is determined using 77 channels between 758 cm^{-1} and 1250 cm^{-1} which we indicate by purple stars in Fig. 1.

In this step, coefficients of six longwave emissivity perturbation functions are solved for, with T_s being held fixed at the value determined from the previously performed skin temperature retrieval step.

3.3 Constituent Profile Retrievals

As in Version-5, constituent profile retrievals are performed in separate subsequent steps, each having its own set of channels and functions. Figure 1 indicates, by stars of different colors, the Version-6 channels used in each of these retrieval steps. The $q(p)$ retrieval (pink stars) uses 41 channels in the spectral ranges 1310 cm^{-1} to 1605 cm^{-1} and 2608 cm^{-1} to 2656 cm^{-1} ; the $O_3(p)$ retrieval (green stars) uses 41 channels between 997 cm^{-1} and 1069 cm^{-1} ; the $CO(p)$ retrieval (gray stars) uses 36 channels between 2181 cm^{-1} and 2221 cm^{-1} ; and the $CH_4(p)$ retrieval (brown stars) uses 58 channels between 1220 cm^{-1} and 1356 cm^{-1} . The Version-6 $q(p)$ retrieval step, including the channels used, is essentially unchanged from that used in Version-5 other than the use of the Neural-Net first guess $q^0(p)$. Some small modifications have been made to the details of the trace gas retrieval steps. Version-6 trace gas retrieval methodology and results are not treated in this paper.

4 Comparison of Quality Controlled Version-6 and Version-6 AO Retrievals with those of Version-5

Our evaluation compares Version-6 and Version-6 AO QC'd products with those of Version-5. In the following sections, we evaluate ocean surface skin temperature T_s , ocean and land surface spectral emissivity, and global temperature profile $T(p)$ and water vapor profile $q(p)$. Our evaluation compares results obtained on nine focus days to collocated 3 hour ECMWF forecasts, which are taken as a measure of truth. The nine focus days are: September 6, 2002;

January 25, 2003; September 29, 2004; August 5, 2005; February 24, 2007; August 10, 2007; May 30, 2010; July 15, 2011; and September 14, 2012. All products have QC flags based on thresholds of error estimates. Both Version-5 and Version-6 use QC flags for the level-2 output products in which QC=0 indicates the best quality products designated for use in a data assimilation application; products flagged with QC=1 are of good quality designated to be included along with those with QC=0 in the generation of gridded level-3 products used for climate research; and products flagged with QC=2 should not be used for any purpose. The Version-6 methodology used to derive error estimates is analogous to that used in Version-5, but their use in the generation of quality flags is somewhat different from that used in Version-5. Details about the generation of error estimates and their use for QC flags are given in Appendix A.

4.1 Ocean Surface Skin Temperature T_s and Surface Spectral Emissivity

The term T_s refers to surface skin temperature over all surfaces. We also refer to values of T_s over non-frozen ocean as Sea Surface Temperature (SST). Figure 2 shows counts of QC'd values of SSTs over the latitude range, 50°N – 50°S, as a function of the difference between T_s and “truth” for the 9-day evaluation period, where “truth” for T_s , and for most other geophysical parameters, is taken from the ECWMF 3-hour forecast field. We show the counts of Version-5 retrievals in red and pink, Version-6 retrievals in dark blue and light blue, and Version-6 AO retrievals in black and gray. The lighter shade of each color shows counts of the best quality T_s retrievals obtained using Data Assimilation error estimate thresholds (QC=0). The darker shade of each color shows counts of both best and good quality T_s retrievals, including cases with QC=0 or 1, where the Climate error estimate thresholds used to obtain QC=1 are looser than those used for QC=0. Ocean T_s retrievals with QC=0 or 1 are the ensemble used over ocean in the generation of

**Surface Skin Temperature Difference from ECMWF
9-Day Average Daytime and Nighttime combined
50 N to 50 S Non-Frozen Ocean**

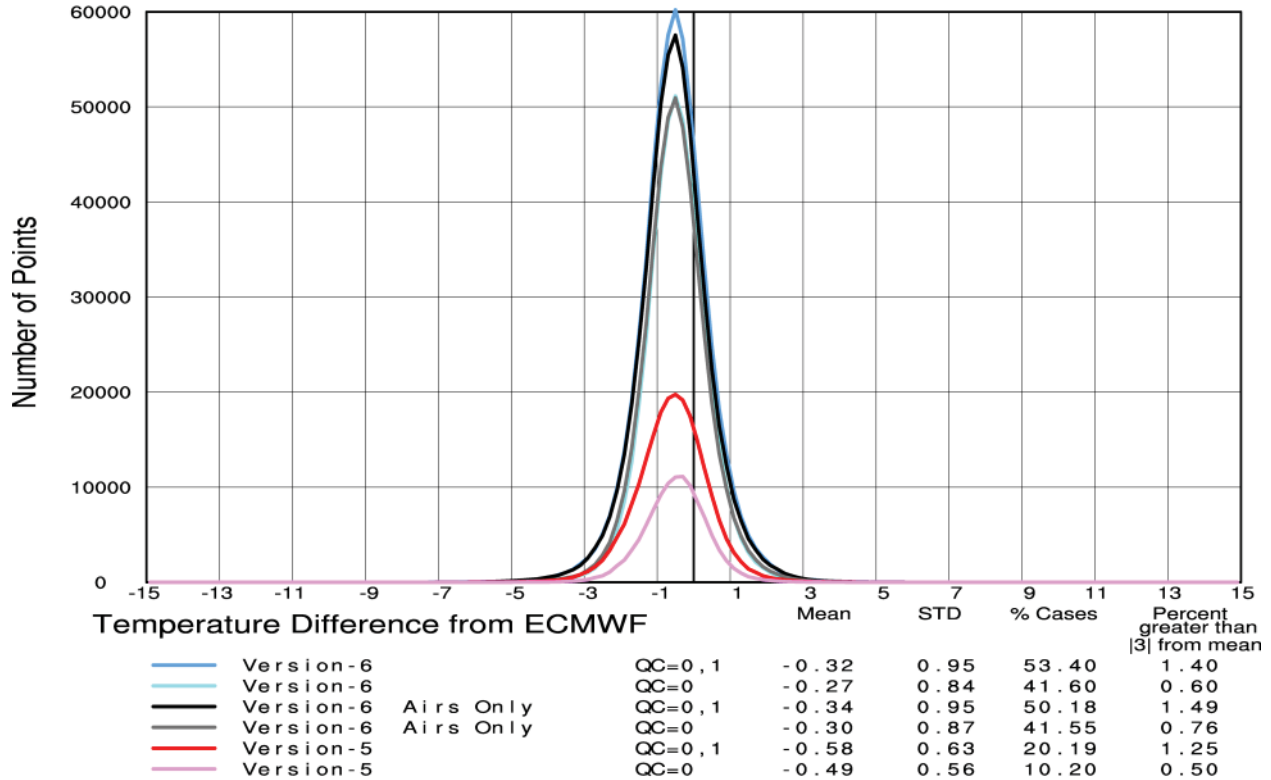


Fig. 2

the level-3 surface skin temperature product used for climate studies. Figure 2 contains statistics for each set of retrievals showing the mean difference from ECMWF, the standard deviation (STD) of the ensemble differences, the percentage of all possible cases included in the QC'd ensemble, and the percentage of all accepted cases with absolute differences from ECMWF of more than 3K from the mean difference, which we refer to as outliers.

Version-6 QC'd retrievals accept considerably more cases than Version-5 and have much lower standard deviations of the errors as well. In both ensembles, the percentage of outliers grows with loosening the QC thresholds, as expected. The percentage of Version-6 outliers with QC=0,1 is somewhat larger than that in Version-5, but the Version-6 yield with QC=0,1 is more

than twice as large as that of Version-5. It is noteworthy that Version-6 retrievals with QC=0 have a much smaller percentage of outliers than do Version-5 retrievals with QC=0,1, along with a substantially higher yield. Statistics of QC'd Version-6 AO retrievals are very similar to those of Version-6.

Figure 3 shows the spatial distribution, over the latitude range 60°N-60°S, of the nine day mean differences of the level-3 oceanic SST products from collocated ECMWF values for both Version-6 and Version-5. The values shown in a given grid box are the average values for that grid box of all cases in which the SST retrieval was accepted using Climate QC either at 1:30 AM or 1:30 PM. The oceanic grid boxes shown in gray indicate grid boxes in which not a single value of Climate QC'd SST occurred for all 18 possible cases (nine days, twice daily). Figure 3 represents the spatial coverage and accuracy of a “pseudo nine day mean” level-3 product.

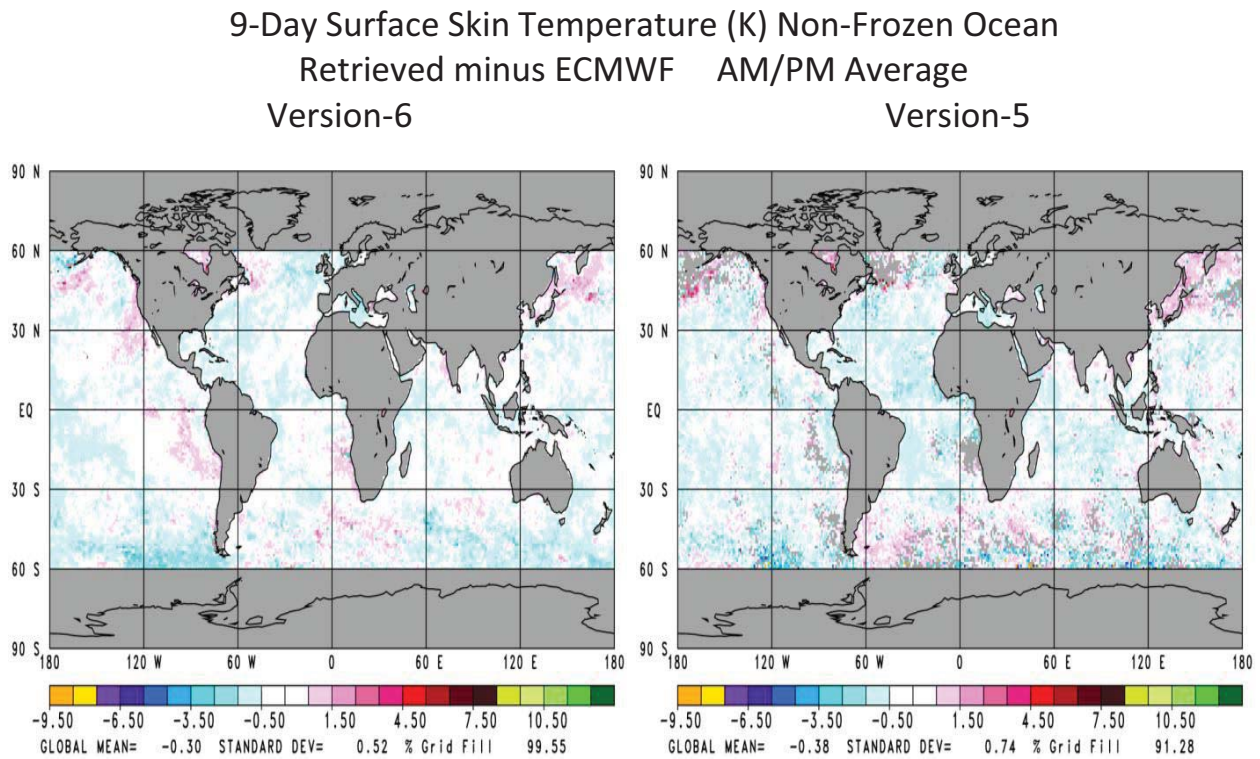


Fig. 3

The results shown in Fig. 3 do not represent those of a typical nine day level-3 product because the nine days used are not consecutive, but Figure 3 provides valuable information nonetheless. The Version-6 “pseudo nine day mean” level-3 product is significantly improved over the Version-5 product in terms of accuracy as compared to ECMWF, and also has almost complete spatial coverage, with 99.55% of possible oceanic grid points covered, while Version-5 has only 91.28% oceanic spatial coverage, and is marked by gaps in areas that had significant cloud cover in each of the 18 time periods included in the nine day mean field.

Figures 4a and 4b show the mean difference of the retrieved ocean surface emissivity from that of the AIRS Science Team ocean surface emissivity model as a function of satellite zenith angle for $\nu = 950 \text{ cm}^{-1}$ and $\nu = 2400 \text{ cm}^{-1}$, and Figs. 4c and 4d show the standard deviations of the retrieved values at a given zenith angle. The two channels shown are in the longwave and shortwave window regions, respectively. In these figures, we show statistics separately for AM orbits in dark colors and PM orbits in light colors. In both the longwave and shortwave window regions, Version-6 (as well as Version-6 AO) retrieved ocean spectral emissivities as a function of satellite zenith angle are very close to the values expected using the AIRS Science Team ocean surface emissivity model. Differences of Version-6 retrieved values of from the ocean emissivity model are much smaller than those of Version-5. Version-5 retrieved values of also showed a large spurious feature during the day in the vicinity of satellite zenith angle -18.24 degrees at both frequencies. This spurious feature occurs at the viewing angle at which maximum sunglint appears in the field of view. In addition to being more accurate in the mean sense, the retrieved values of are much more stable in Version-6 compared to those of Version-5, as evidenced by the much lower standard deviations of their values as shown in

Ocean Surface Emissivity vs. Zenith Angle

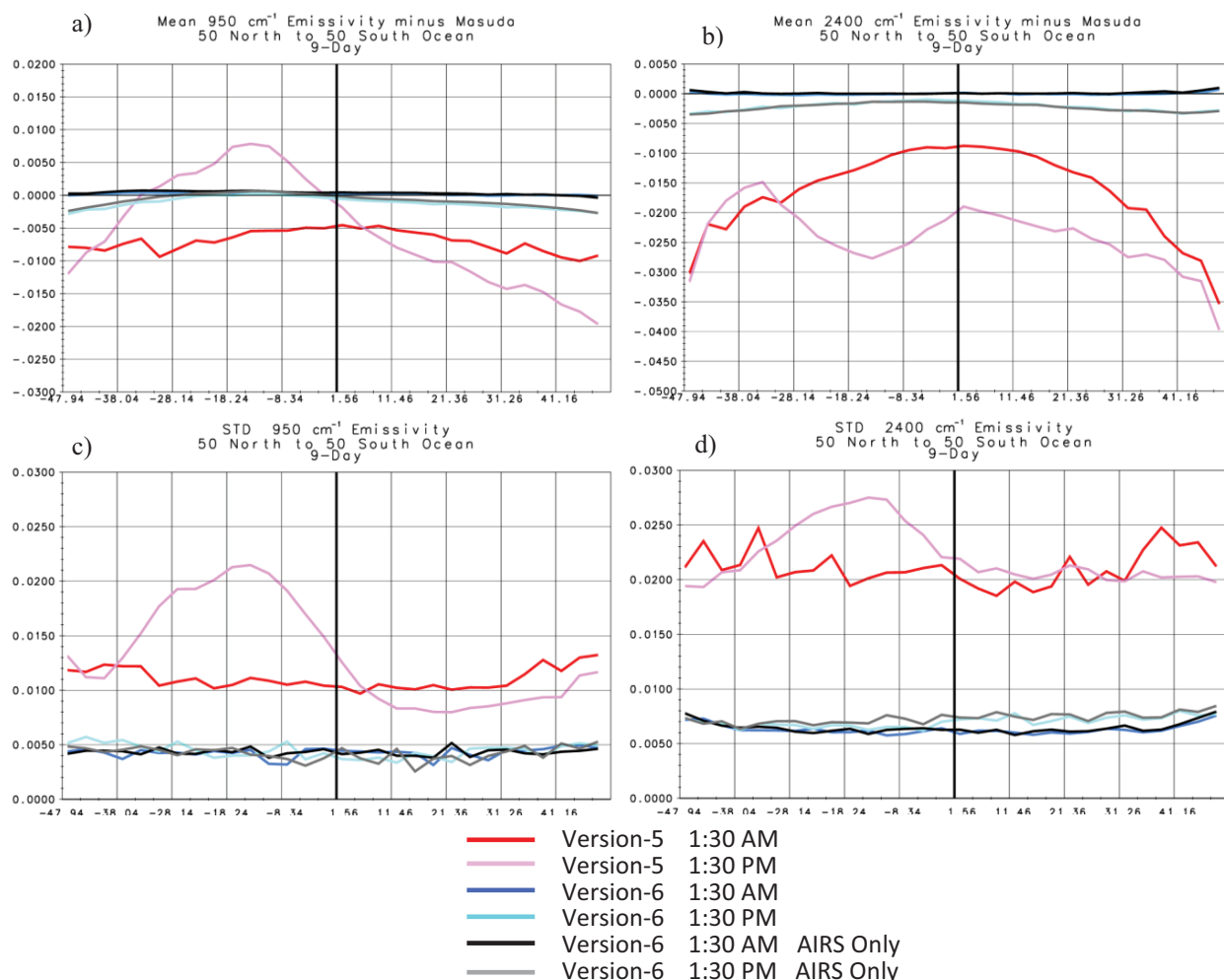


Fig. 4

Figs. 4c and 4d. There is no appreciable difference between Version-6 and Version-6 AO results related to retrieved ocean values of .

Figures 4a and 4b show that daytime and nighttime Version-6 retrieved values of ocean surface emissivity are not only close to those of the ocean emissivity model, which is a good measure of truth, but they are also very close to each other, as expected. Over land, surface spectral emissivity values change rapidly in space, and time as well, as a result of variations in ground cover, such as vegetation, rock and soil types, and even snow cover. At a given location

and day, these values should not change appreciably from day to night, however. Figure 5 shows the nine-day mean 1:30 AM/1:30 PM differences of retrieved values of ϵ at 950 cm^{-1} and 2400 cm^{-1} over land obtained using the Version-6 and Version-5 retrieval systems. As in the case of ocean, day/night differences of Version-6 retrieved land surface emissivity are much smaller than those of Version-5, as they should be.

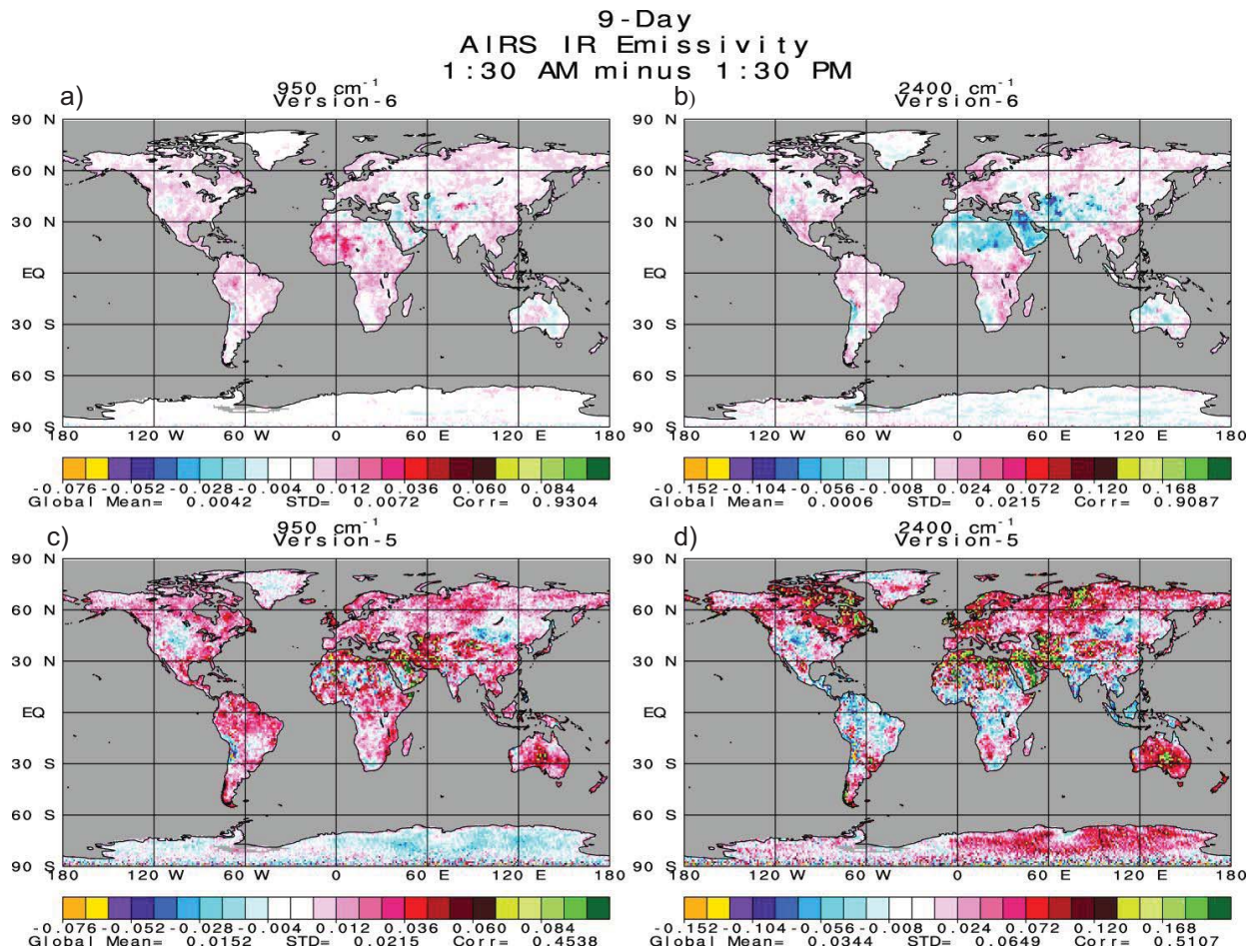


Fig. 5

4.2 $T(p)$ Retrieval Accuracy as a Function of Yield

The fundamentals of the methodology used in Version-6 to retrieve temperature profile $T(p)$ from AIRS cloud cleared radiances are basically the same as those used in Version-5. Figure 6

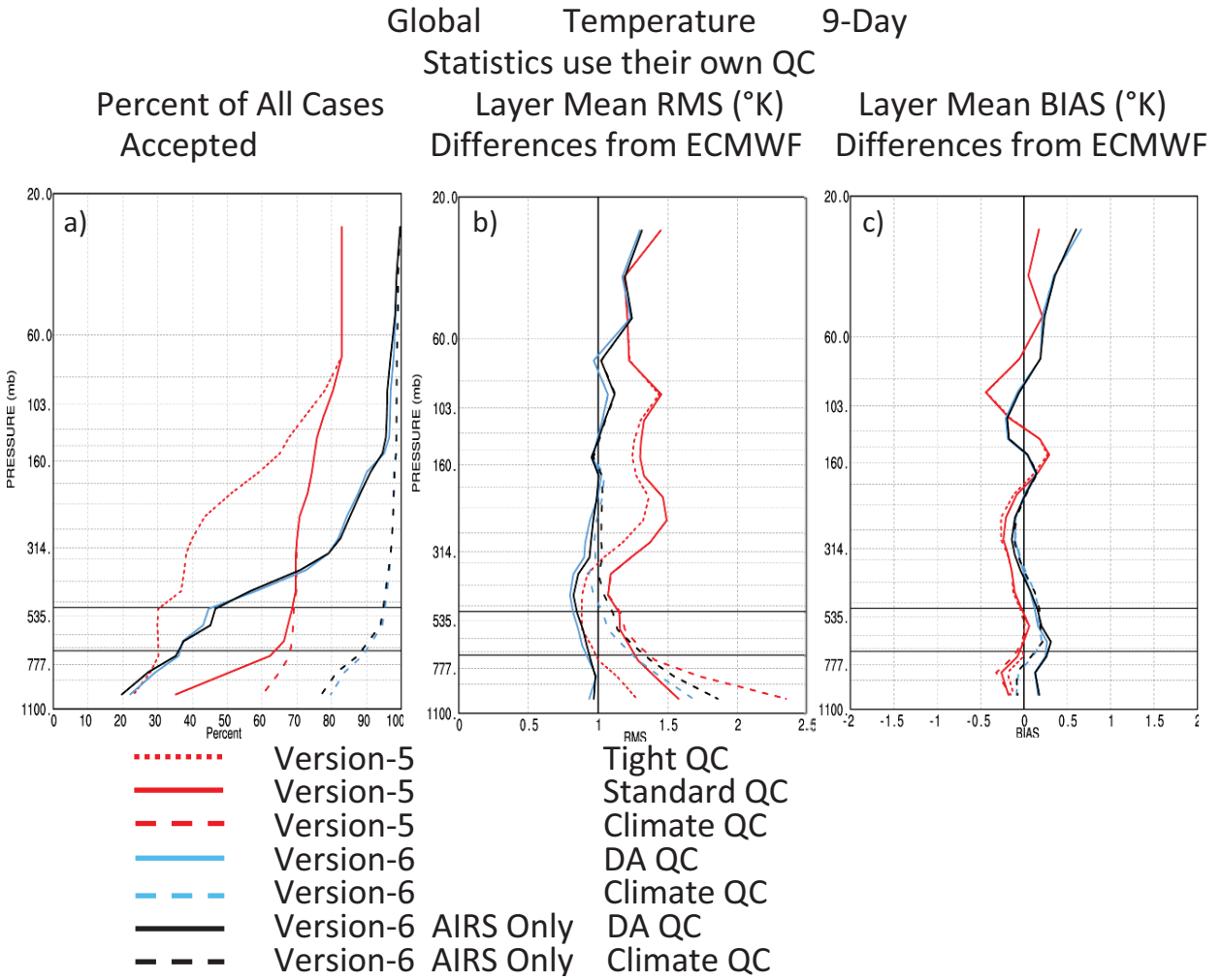


Fig. 6

shows statistics of the differences of QC'd Version-5 and Version-6 $T(p)$ retrievals from collocated ECMWF truth for a global ensemble of cases taken over the nine focus days. Panel (a) shows the percentage of QC'd cases accepted as a function of height, panel (b) shows RMS differences of 1 km layer mean temperatures from collocated ECMWF "truth", and panel (c) shows biases of QC'd 1 km layer mean differences from ECMWF. Statistics are shown for seven sets of results. We show in red the results for Version-5 retrievals using three different QC

procedures; in blue results for Version-6 retrievals using two different QC procedures; and in black results for Version-6 AO retrievals using analogous QC procedures to those of Version-6.

QC procedures used in both Version-5 and Version-6 designate two characteristic pressures for each temperature profile, p_{best} and p_{good} . These pressures are computed using thresholds of temperature profile error estimates $\delta T(p)$. Appendix A describes the manner in which $\delta T(p)$ is computed and how it is used for QC purposes. Version-5 had only one set of $T(p)$ QC error estimate thresholds, called Standard Thresholds, which were used to define Version-5 values of p_{best} , down to which $T(p)$ retrievals were considered to be of highest quality. The Version-5 Standard Thresholds were chosen such that if one utilized only $T(p)$ retrievals down to p_{best} for each case, this procedure would provide a middle ground of keeping retrievals with highest accuracy, which would be optimal for data assimilation (DA) purposes on the one hand, and keeping retrievals with the highest yield (best spatial coverage), optimal for climate purposes on the other hand. Experience using Version-5 products showed that Standard QC Thresholds were optimal for neither purpose. For example, data assimilation experiments assimilating Version-5 retrievals down to a value of p_{best} defined using a tighter set of thresholds than found in the official Version-5 system, referred to as Tight Thresholds⁸, resulted in significantly improved forecasts compared to assimilation of $T(p)$ retrievals down to values of p_{best} computed using the looser Standard QC thresholds. The dotted red lines in Fig. 6 show acceptance yield and accuracy of Version-5 retrievals down to p_{best} as defined using the Tight QC thresholds (not officially part of Version-5). The solid red lines in Fig. 6 show equivalent statistics for the ensemble of Version-5 retrievals down to p_{best} as computed using the Standard Thresholds. The global yield of cases in which p_{best} is equal to the surface pressure p_{surf} , as defined using Standard Thresholds, is shown in Fig. 6a to be about 35%. Utilization of an ensemble of

retrievals with such a low yield would not be adequate for the generation of level-3 $T(p)$ products with reasonable spatial coverage near the surface. The spatial coverage near the surface of such an ensemble of cases is particularly poor over land and sea-ice. In order to be able to generate level-3 products with reasonable spatial coverage in Version-5, an additional case by case characteristic pressure, p_{good} , was defined in an ad-hoc manner over land and sea-ice for use in the generation of level-3 products. If p_{best} was at least 300 mb over these domains, p_{good} was set to be equal to the surface pressure p_{surf} . Otherwise, p_{good} was set equal to p_{best} . Version-5 level-3 products for $T(p)$ at a given pressure p were generated using all cases for which p was $\leq p_{good}$. Over non-frozen ocean, there was no need to include additional cases in the generation of a level-3 product with good spatial coverage, and p_{good} over non-frozen ocean was always set equal to p_{best} . The dashed red lines in Fig. 6 show statistics for Version-5 retrievals which are included down to p_{good} , i.e., statistics for the ensemble of cases used in the generation of the Version-5 level-3 $T(p)$ products. Global yield of Version-5 cases down to p_{good} at the surface is about 60%, and global mean RMS error of Version-5 cases down to p_{good} is about 2.7K near the surface.

Having learned from the experience with Version-5 QC methodology based on use of a single set of $T(p)$ thresholds for both data assimilation and climate applications, Version-6 defines p_{best} and p_{good} independently of each other based on use of two different sets of QC thresholds: a tight set of DA $T(p)$ thresholds, optimized for data assimilation purposes (QC=0), was designed to derive p_{best} , and a substantially looser set of $T(p)$ thresholds optimal for climate purposes (QC=1) was used to derive p_{good} . The solid blue and black lines in Fig. 6 show statistics for Version-6 and Version-6 AO results respectively using their appropriate sets of DA QC thresholds, including all cases down to p_{best} , and the blue and black dashed lines show results

using the appropriate climate thresholds, including all cases down to p_{good} . As in Version-5, Version-6 and Version-6 AO level-3 gridded products utilize all cases passing climate QC, that is, all cases down to p_{good} .

In Version-5, all retrievals were either accepted or rejected above 70 mb based on use of different types of tests, even before applying the error estimate based QC procedures³. One of the tests that disqualified the entire temperature profile, and flagged the entire profile with QC=2 (do not use), is that the retrieved cloud fraction was over 90%. Roughly 83% of Version-5 retrievals passed the initial screening procedure, with none of them occurring under near overcast conditions. Version-5 retrievals with Tight QC have considerably lower yield than those with Standard QC below 200 mb, with correspondingly smaller RMS errors on the order of 1K beneath 300 mb. The ensemble of Version-5 retrievals used to generate level-3 $T(p)$ products (dashed red line) differs from that of those accepted down to p_{best} below 300 mb. The yield near the surface is roughly 60%, which is better for the generation of level-3 products but the RMS error for this larger ensemble of cases with QC=0 or QC=1 is much larger near the surface than those with QC=0. While RMS errors of retrievals increased with increasing yield, there is no appreciable difference in Version-5 bias errors, compared to ECMWF, found using any of the three ensembles of cases.

Version-6 does not apply any test which eliminates the entire temperature profile, other than the requirement that the retrieval runs to completion. Version-6 retrievals using DA thresholds (QC=0) have a yield much higher than those passing Version-5 Tight thresholds down to about 700 mb, and have RMS errors less than 1K at all levels, which has been found to be optimal for data assimilation purposes⁸. Among other benefits from the perspective of data assimilation is that Version-6 will allow for the assimilation of AIRS temperature products above the clouds,

both in storms, as well as under overcast conditions in general. The yield of Version-6 retrievals with Climate QC (QC=0,1) is extremely high throughout the atmosphere, with a value of about 80% at the surface. Achievement of this very high yield is extremely valuable in the generation of more representative Version-6 level-3 products used for climate studies. RMS errors of Version-6 retrievals with Climate QC are better than, or comparable to, those of Version-5 with Standard QC down to the surface, and significantly better than that of the ensemble of Version-5 retrievals used to generate level-3 products. Results for Version-6 AO using either QC procedure are roughly comparable to those of Version-6, but with slightly lower yields near the surface.

Version-6 retrievals are more accurate than those of Version-5 for a number of reasons. A substantial improvement in Version-6 lower tropospheric temperature profiles results from the improvements made to surface skin temperatures and surface emissivities, especially over land. Version-6 also benefits considerably from the use of a Neural-Net first guess, which is not only more accurate than the regression-based guess used in Version-5, but also degrades much more slowly with increasing cloud cover than does the regression guess. Figures 7a and 7b compare RMS errors of QC'd Version-6 and Version-5 retrievals with those of their first guesses. The solid and dashed blue and red lines shown in Fig. 7 are identical to those in Fig. 6b. The RMS errors of the first guesses are shown by light blue lines for Version-6 and pink lines for Version-5. Figure 7a shows that the Version-6 retrievals improve on the Neural-Net guess at all pressures greater than about 150 mb, especially for the easier ensemble of cases accepted using DA thresholds. Version-6 retrievals are slightly poorer than their first guess above 60 mb. Figure 7b shows that essentially the same relative result holds for Version-5, though in Version-5, the retrievals improve on their first guess at all levels. In addition, unlike for Version-6, the improvement over the first guess is greatest in the mid troposphere.

Layer Mean RMS Temperature ($^{\circ}\text{K}$) Differences from ECMWF
9-Day Global Average

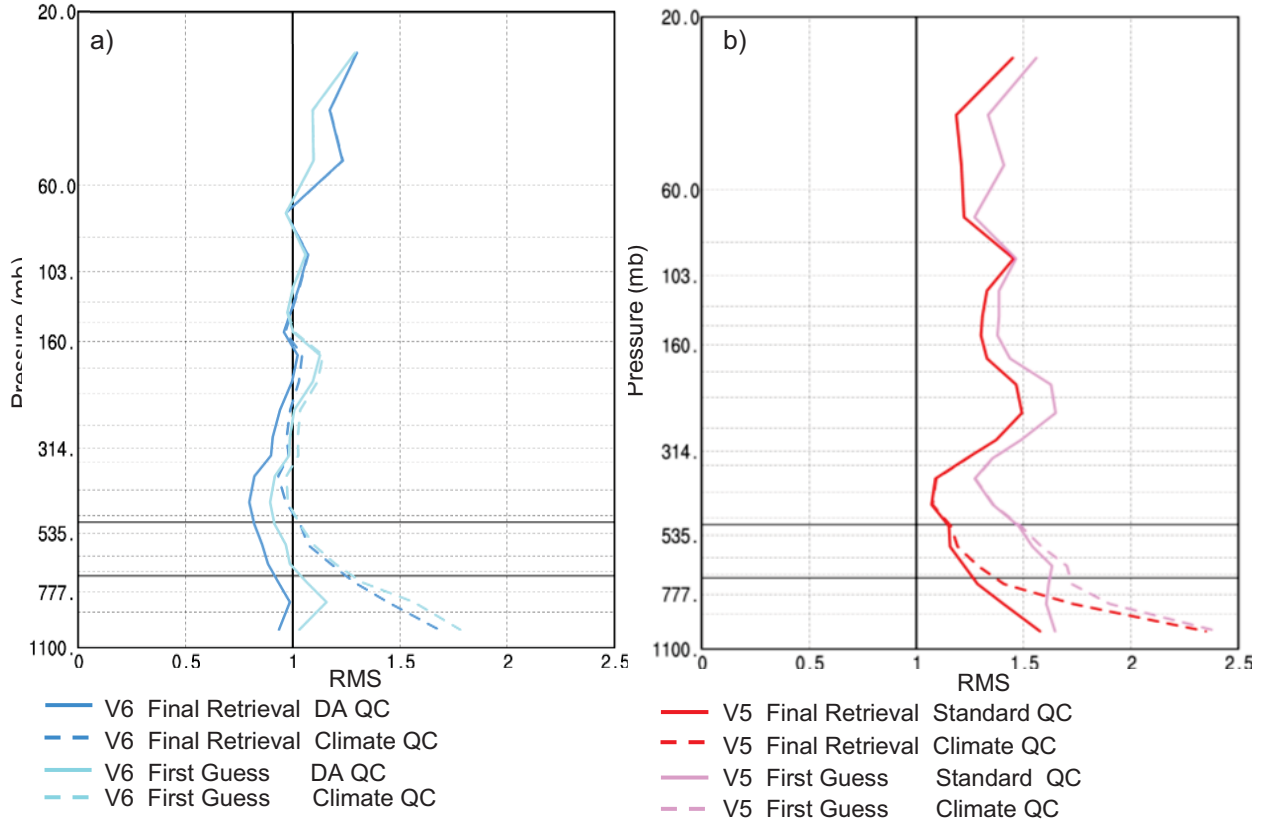


Fig. 7

Figure 8a shows % yields of Version-5 and Version-6 retrievals, accepted using Version-5 Standard QC and Version-6 Climate QC respectively, as a function of retrieved cloud fraction at three mid-lower tropospheric pressures, and Fig. 8b shows the RMS $T(p)$ errors over three corresponding 1 km layers. Version-6 Climate QC yields are much higher than those of Version-5 at all cloud fractions, especially at larger cloud fractions. Version-6 RMS errors over these larger ensembles of cases for all cloud fractions are also considerably better than those of the smaller Version-5 ensembles. The fact that Version-6 retrievals remain accurate and improve over the Neural-Net first guess at larger cloud fractions indicates that the Version-6 cloud cleared radiances are accurate as well under more difficult cloud conditions.

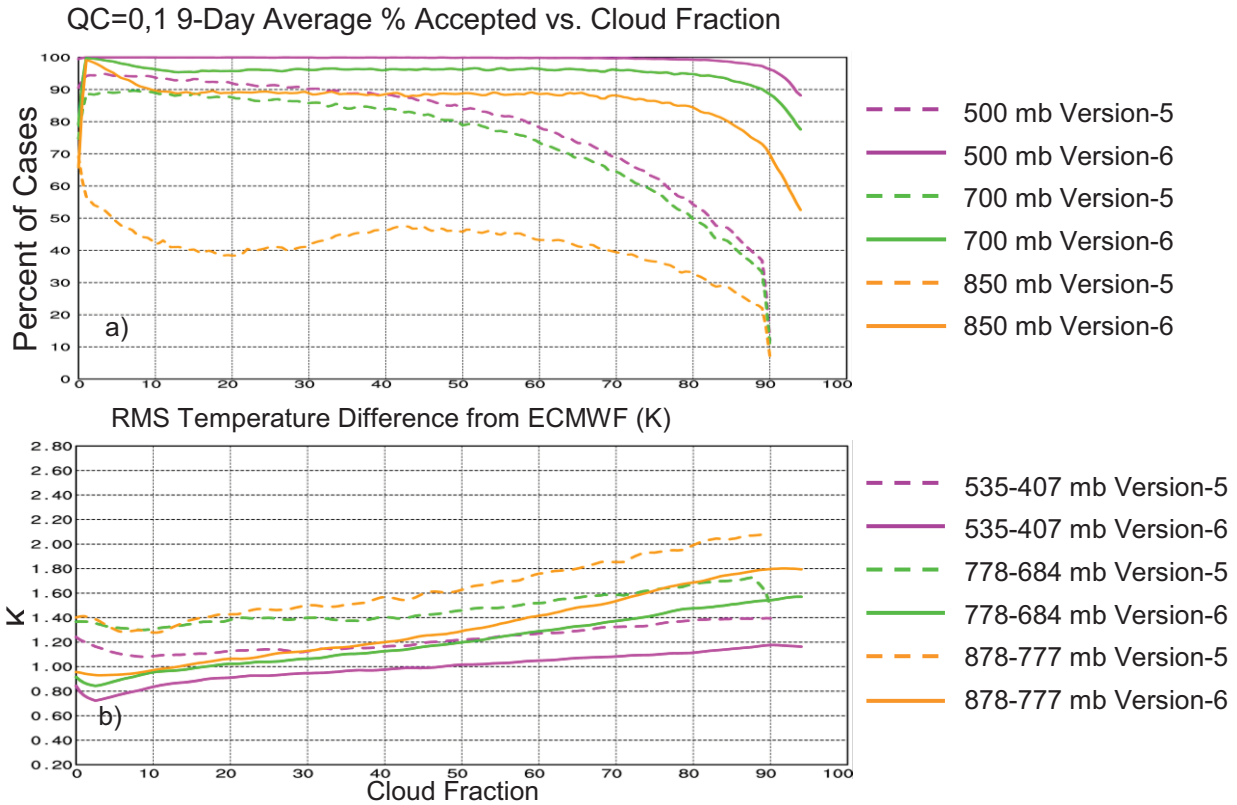


Fig. 8

4.2.1 Yield trends and spurious bias trends of $T(p)$

Our research using Version-5 retrieved products indicated that QC'd Version-5 values of $T(p)$ had a large negative yield trend, as well as spurious bias trends when compared to collocated ECMWF values of $T(p)$. A prime consideration in the finalization of Version-6 was to alleviate these negative yield trends and spurious bias trends as much as possible. Figure 9 shows yield trends and temperature bias trends of Version-5 retrievals using Standard QC, and both Version-6 and Version-6 AO retrievals using Climate QC, as evaluated over the nine days used in all other figures. Figure 9a shows that the % yield of accepted Version-5 retrievals was decreasing over time (negative yield trend), and Figure 9b shows that Version-5 retrievals had substantial negative spurious temperature bias trends in the troposphere. A substantial part of the negative

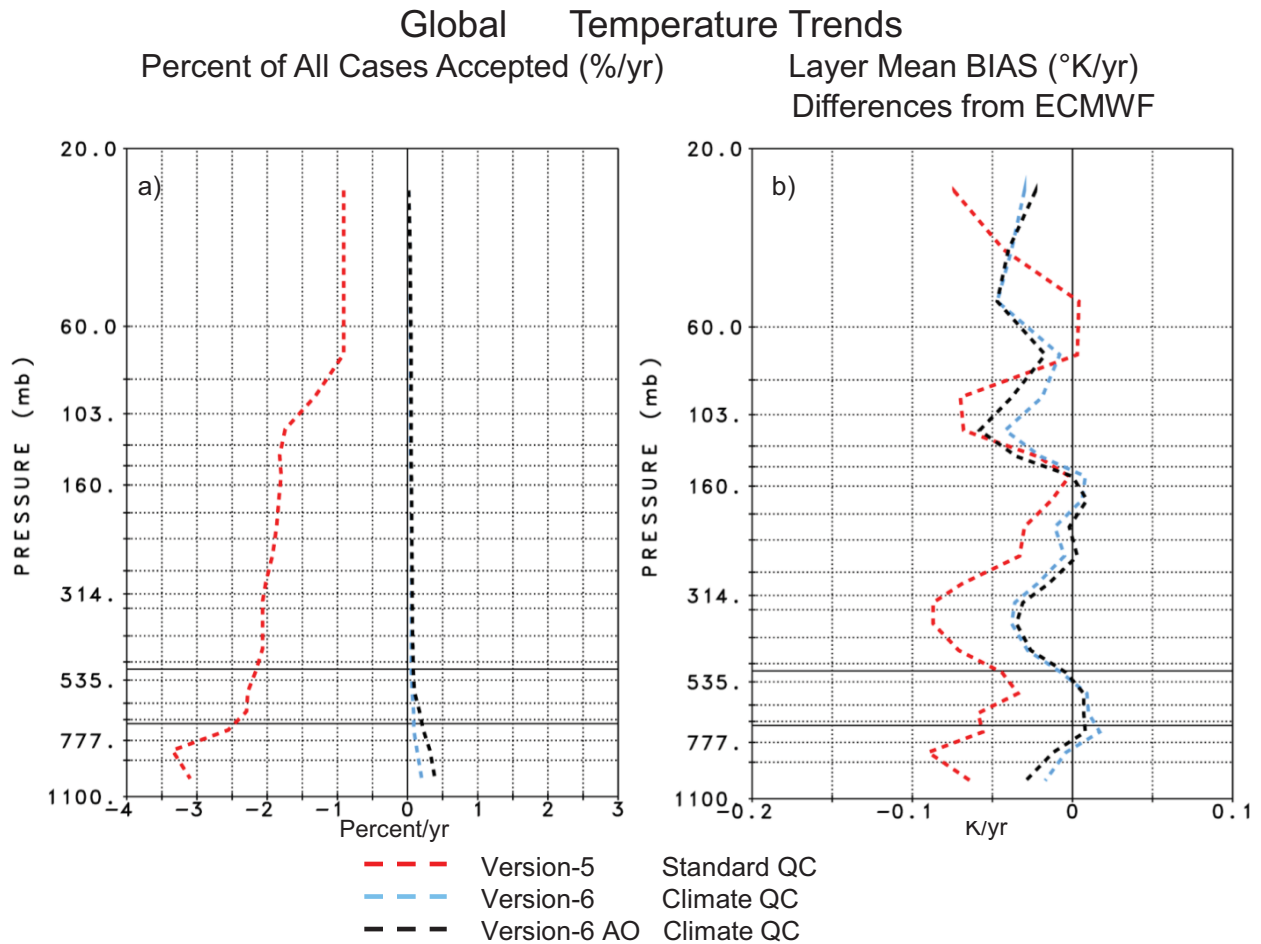


Fig. 9

yield trend was due to a significant degradation of the noise characteristics of AMSU-5. Version-6 contains modifications which alleviated these problems, one of which is that Version-6 no longer uses AMSU-5 at all. Other factors also contributed to the spurious temperature bias trends found in Version-5, and these were also corrected in Version-6.

Figure 9 shows that Version-6 has eliminated the substantial negative tropospheric temperature profile yield trends, on the order of 2% per year, which were found in Version-5. In addition, the Version-6 negative $T(p)$ bias trends beneath 500 mb are much smaller than those of Version-5, which were as large as -0.08K/yr . There is no appreciable difference between the yield or bias trend results obtained for Version-6 and Version-6 AO.

4.3 Retrieval Accuracy of $q(p)$

The details of the $q(p)$ retrieval step are essentially unchanged from what was done in the $q(p)$ retrieval step both in Version-5 and in Version-4. Version-7 will address further improvements to be made to the $q(p)$ retrieval algorithm. Nevertheless, Version-6 retrieved values of $q(p)$ are improved over those of Version-5 as a result of the same factors that led to improved Version-6 values of $T(p)$ as compared to Version-5: 1) improved surface skin temperatures and spectral emissivities; 2) an improved first guess $q^0(p)$ provided by the Neural-Net start-up system; and 3) improved clear column radiances . Version-6 retrieved values of $q(p)$ also benefit from improved values of $T(p)$ that are used as input to the $q(p)$ retrieval step.

Figure 10 shows analogous results to those of Fig. 6 comparing QC'd 1 km layer precipitable water to that of collocated values of ECMWF. We show results only up to 200 mb, above which water vapor retrievals are considered to be of minimal validity and are not included in the AIRS Science Team Standard Product data set. The relative results comparing Version-5 and Version-6 $q(p)$ retrievals are analogous to those found for $T(p)$. Version-6 $q(p)$ retrievals with both DA and Climate QC are considerably improved over those of Version-5 in the lower troposphere. This improvement in the lower troposphere is at least partially a result of the improved values of T_s and ε_v in Version-6 compared to Version-5. As with $T(p)$, Version-6 $q(p)$ retrievals with Climate QC are unbiased, have high accuracy, and contain almost complete spatial coverage. Globally, Version-6 AO $q(p)$ retrievals are slightly less accurate than those of Version-6 near the surface. This difference between results of Version-6 and Version-6 AO occurs primarily over the ocean and is a result of the benefit over ocean of the 22 GHz and 31 GHz channels of AMSU-A, which are not included in the AIRS Only retrieval procedure.

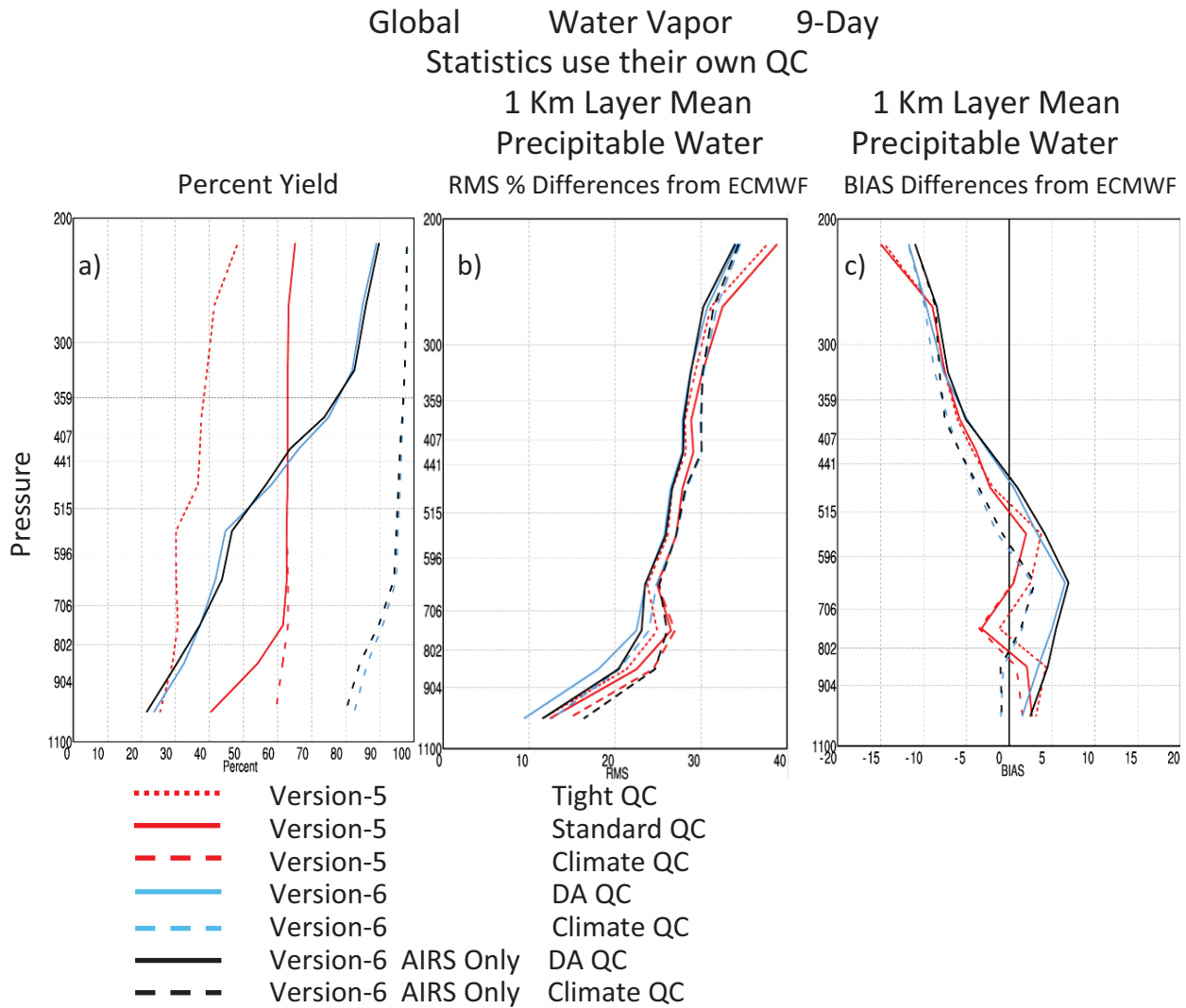


Fig. 10

Figure 11 shows the spatial distribution of the “pseudo-level-3” nine-day mean field of accepted cases of total precipitable water, W_{TOT} , flagged to be of Climate quality (QC=0,1). The statistics shown for Version-6 and Version-5 represent the area weighted global mean difference and spatial standard deviation of the gridded level-3 values of W_{TOT} from the collocated ECMWF value of W_{TOT} . Statistically, the Version-6 “pseudo nine-day mean” level-3 values of W_{TOT} are considerably more accurate than those of Version-5, especially over land.

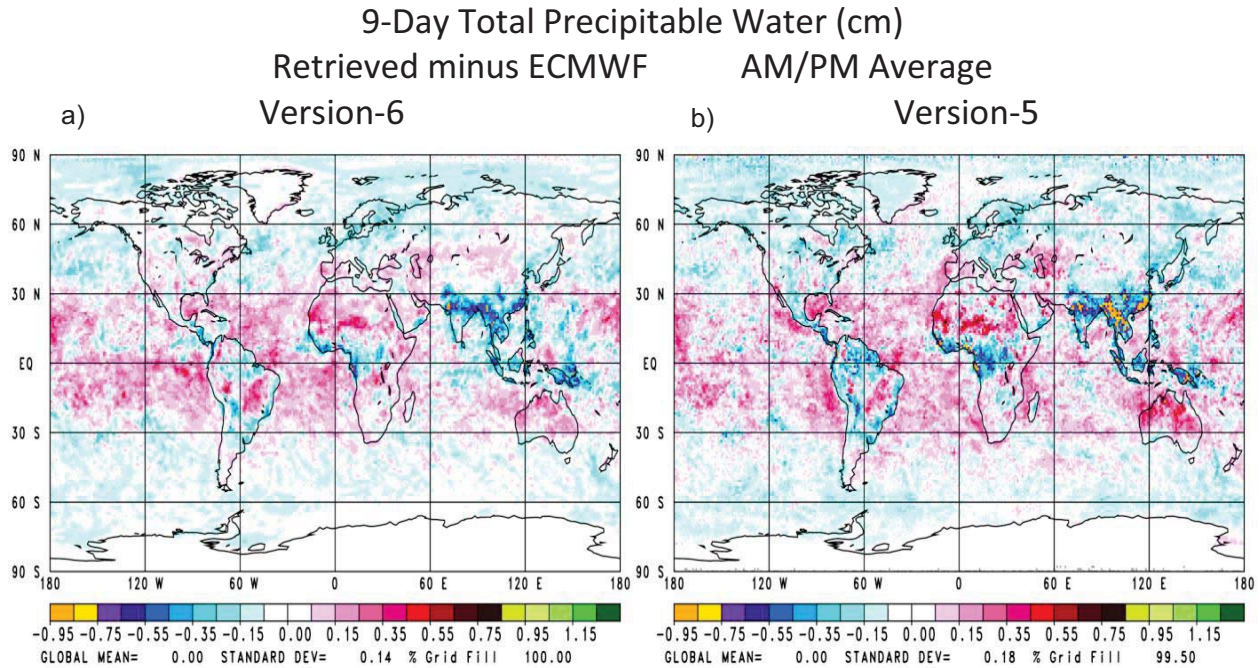


Fig. 11

In Version-6, W_{TOT} is flagged to be of Climate quality if the entire water vapor profile has best (QC=0) or good (QC=1) quality down to the surface. This same test is also applied to generate QC flags for 1) surface air temperature; 2) clear sky OLR; 3) O₃, CH₄, and CO profiles; and 4) surface skin temperature over land and frozen ocean. Version-5 used a different procedure to accept those values of W_{TOT} to be used in the generation of the level-3 product.

4.4 Comparison of Version-6 and Version-5 Retrieved Values of Cloud Fraction and Cloud Top Pressure

The procedure used to derive cloud fraction and cloud top pressure in Version-6 is similar to that used in Version-5^{1,3}, but has a number of significant improvements. The radiatively effective cloud fraction at frequency ν , is given by the product of α , the geometric fractional cloud cover of an AIRS FOV as seen from above, and ϵ , the cloud spectral emissivity. The AIRS Science Team cloud parameter retrieval methodology determines only the product of these two

terms, τ_{λ} , along with a corresponding cloud top pressure p_c , for each of up to two layers of clouds in a given scene^{1,3}. A basic assumption of the cloud retrieval methodology used in both Version-5 and Version-6 is that the clouds are gray, that is, τ_{λ} is independent of frequency. Version-5 simultaneously derived 20 parameters for each AIRS FOR, nine effective cloud fractions f_{λ} and f_{λ}^* one pair for each AIRS FOV ℓ contained within the AMSU FOR, along with two cloud top pressures p_{c1} and p_{c2} , and p_{c1} considered to be representative of the pressures of each of the two layers of clouds covering the entire AIRS FOR. In Version-6, the cloud parameter retrieval step is performed separately for each AIRS FOV ℓ to determine four parameters, f_{λ} , f_{λ}^* , p_{c1} , and p_{c2} , in each FOV. A total radiatively effective cloud fraction for the entire FOR, f_{FOR} , is computed as the average cloud fraction according to

$$(4)$$

and an effective cloud top pressure for the entire FOR is computed as the weighted average of all 18 values of p_c in the FOR

$$(5)$$

as was also done in Version-5. The Version-6 level-2 product contains individual values of f_{λ} and f_{λ}^* for each AIRS FOV, as well as the single FOR heritage values f_{FOR} and p_{c1} defined according to Equations 4 and 5.

Cloud parameters in an AIRS FOV are derived such that channel radiances computed using these cloud parameters τ_{λ} and p_c , where X is a state vector for the FOV, best match the observed radiances R_{λ} in that FOV for the ensemble of cloud retrieval channels i . The i channels used to determine cloud fraction and cloud top pressure are the same as those used in the cloud clearing step and are shown by yellow stars in Fig. 1. The state vector X used to derive

cloud parameters in an AIRS FOV is the geophysical state retrieved for the entire AIRS FOR containing the nine FOV's.

In Version-5, the state vector X used to derive values of τ and τ_{eff} in a FOR was the retrieved state used in the final cloud clearing step for those cases in which a successful combined AIRS/AMSU retrieval was performed. In the roughly 27% of the cases in which the AIRS/AMSU retrieval was rejected (see Fig. 6), the state X used to derive cloud parameters was the so called “fallback state” that was obtained from a previously performed “AMSU Only” retrieval step³. Cloud parameters retrieved using the fallback state vector X were flagged as QC=1, and those retrieved using the final retrieval state vector X were flagged as QC=0. Under some conditions, the cloud parameter retrieval step was not able to complete successfully, and clouds retrieved for those cases were flagged as QC=2 in Version-5.

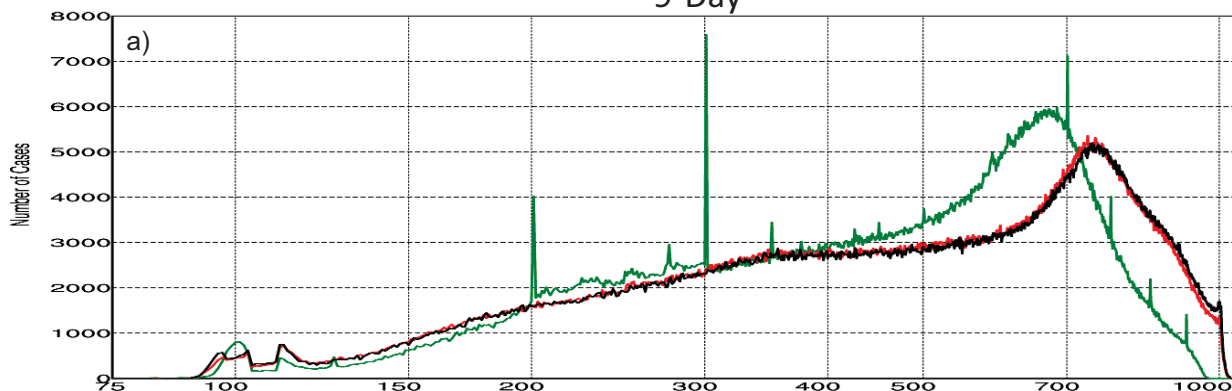
In Version-6, successful AIRS/AMSU retrievals are performed under essentially all conditions so there is no need to use X derived from a microwave “fallback state”. Nevertheless, Version-6 does resort to use of X derived from a partial “fallback state” under some circumstances in which part of the retrieved state X is known to be of poor quality and a better alternative is available. In particular, values of T_s retrieved under either near overcast or overcast conditions over ocean can be spuriously very low. These values of T_s will in general be flagged as bad, with QC=2, meaning they are not used in the generation of the level-3 T_s product. Associated values of τ retrieved under these conditions will also be poor, and are also flagged to be of poor quality. Nevertheless, some value for T_s and τ must be included in the state vector X used to derive the cloud product. We have found that the initial guess T_s^o coming from the Neural-Net start-up procedure gives reasonable values over non-frozen ocean even for very cloudy cases. Therefore, over ocean, if $|T_s - T_s^o| > 5\text{K}$, we assume the retrieved values τ and T_s

are in error, and replace T_s and τ in the retrieved state vector X by T_s^o , and τ^o while retaining the remainder of the retrieved state vector X when computing cloud parameters. We have found that Neural-Net values of T_s^o over land or ice are not of sufficiently good quality for use in the generation of cloud parameters, so this test and replacement procedure is done only over open ocean. As in Version-5, cloud parameters retrieved in such “fallback” cases are flagged as QC=1. Cloud parameters retrieved under almost all other cases, which represent the vast majority of the cases, are flagged as QC=0. Under the extremely rare conditions in which the final cloud parameter retrieval step does not complete successfully, cloud parameters are flagged as QC=2 as was done in Version-5.

A complication in the cloud parameter retrieval methodology is that the best least squares fit may result from a cloud parameter solution which lies in a region which is unphysical. In particular, we do not allow retrieved cloud fractions to be less than zero or greater than 100%, nor do we allow cloud top pressures to be very close to the surface or above the tropopause. Because of the way these constraints were handled in Version-5, many cloud retrievals in Version-5 failed to converge properly. We made numerous small enhancements in the details of the cloud parameter retrieval step in Version-6 which alleviated this problem.

Figure 12a shows the number of cases in which a non-zero cloud fraction $\alpha\epsilon$ was retrieved as a function of cloud top pressure p_c for Version-5, Version-6, and also for Version-6 AO. Two features are readily apparent from Fig. 12a; the distributions of the number of cases obtained as a function of retrieved cloud top pressure are essentially identical in Version-6 and Version-6 AO; and both are substantially different from that of Version-5. Version-5 has spikes in the number of cases retrieved at select pressures, such as 200 mb, 300 mb, 350 mb, 750 mb, 850 mb, and 950 mb, which result from the cloud retrieval algorithm’s inability to converge properly in those

Number of Cases for Average Cloud Top Pressure
Daytime/Nighttime Average
9-Day



Total Cloud Fraction as a Function of Average Cloud Pressure
Daytime/Nighttime Average
9-Day

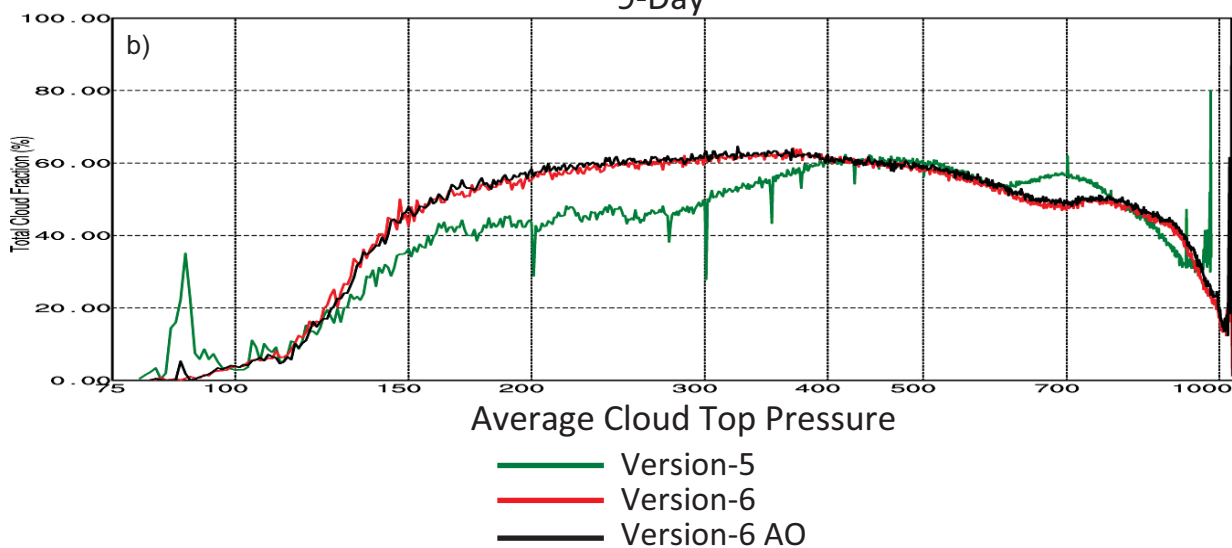


Fig. 12

cases. Such features are not observed in either Version-6 or Version-6 AO. Even more significant is the shift to higher pressures in the peak of the occurrence of low clouds in Version-6 as compared to Version-5. This difference near 1000 mb is in part due to the constraint used in Version-5 that p_c must be at least 50 mb above the surface, while in Version-6, p_c was allowed to go down to 10 mb above the surface. The large shift in the peak in the number of clouds retrieved in Version-5 as a function of cloud top pressure, from roughly 650 mb, to

about 750 mb in Version-6, is a combined result of changes not only in the cloud parameter retrieval step, but also in the state vector X used in Version-5 compared to that used in Version-6, which does not use a “microwave only” fallback retrieval state.

Figure 12b shows analogous plots to those shown in Fig. 12a, but shows the average cloud fraction $\alpha\epsilon$ found for each cloud top pressure p_c . Cloud fractions in Version-6 and Version-6 AO are again very close to each other and differ significantly at some cloud top pressures from those of Version-5. Version-6 has more clouds than Version-5 between 130 mb and 400 mb. On the other hand, Version-6 has less clouds than Version-5 between about 600 mb and 750 mb, which correspond to pressures at which the maximum numbers of cloud parameter retrievals occurred in Version-5. Figure 12b shows spikes in the retrieved cloud fraction in Version-5 at the same pressures in which they occurred in Fig. 12a. These Version-5 spikes in Fig. 12b are negative at pressures lower than 500 mb, indicative of the fact that the spurious cloud retrievals occurring at these discrete pressures had low, probably near zero, cloud fractions. On the other hand, these spikes in Fig. 12b for Version-5 were positive at pressures 700 mb and greater, indicative that these spurious cases had large cloud fractions, most likely close to 100%. Version-5 also had a somewhat disconcerting peak near 90 mb in Fig. 12b, but Fig. 12a shows that there were very few such cases.

Figure 13 shows the spatial distributions of values of cloud fraction $\alpha\epsilon$ and cloud top pressure p_c for the daytime and nighttime orbits on September 29, 2004 as retrieved using Version-5 and Version-6. These plots depict both $\alpha\epsilon$ and p_c at the same time. There are seven different color scales used for different intervals of p_c , as indicated on the figures. Reds, violets, and purples indicate high clouds, blues and greens indicate mid-level clouds, and oranges and yellows indicate low clouds. Within each color scale, darker colors indicate larger fractional cloud cover,

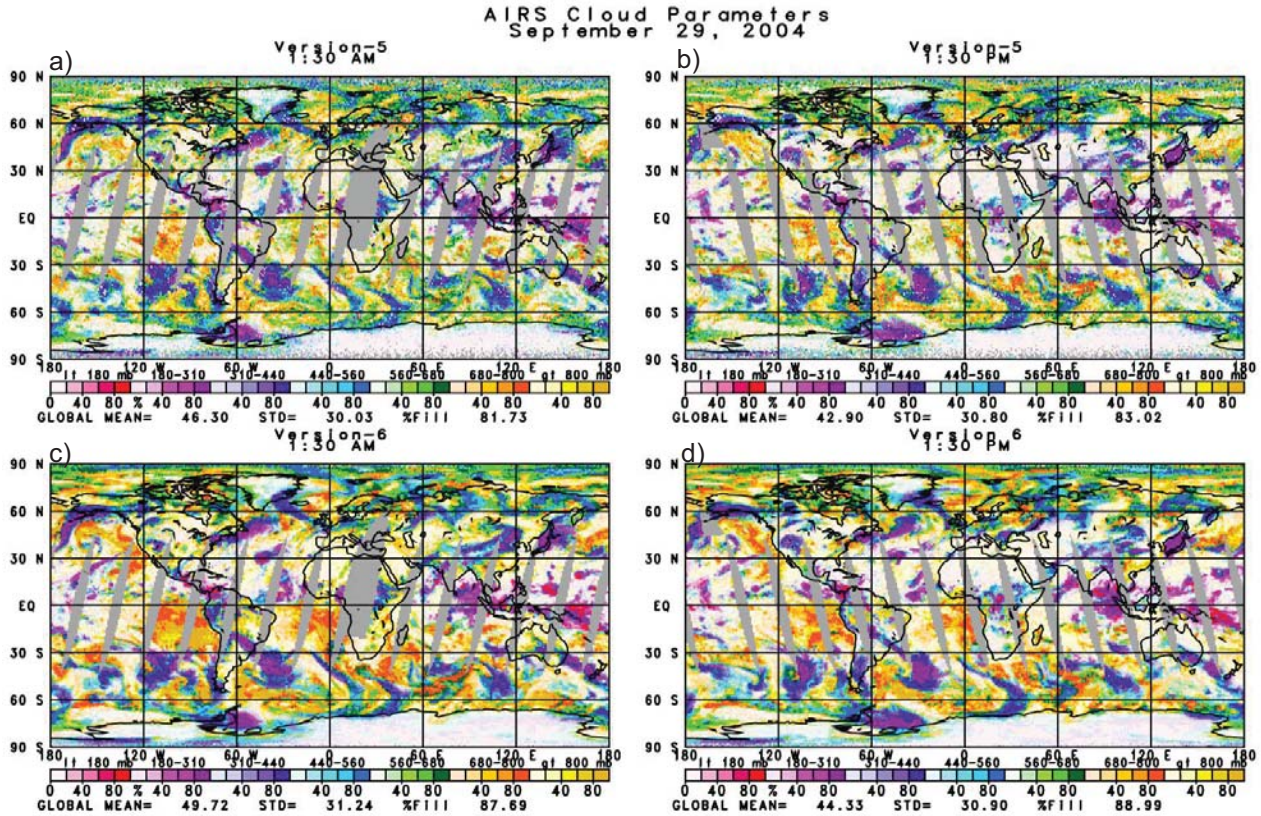


Fig. 13

and paler colors indicate lower fraction cloud cover. While the basic cloud patterns are the same in Version-6 and Version-5, the cloud features are much more coherent, and the colors are darker, in Version-6. Of particular significance are the coherent areas of dark orange, depicting extensive cloud cover with cloud top pressures between 680 mb and 800 mb, found in Version-6 that are at best muted in Version-5, especially at 1:30 AM. This finding is indicative of the better ability to derive the existence of stratus clouds over ocean in Version-6 as compared to Version-5. Particularly noteworthy is the region in the vicinity of 10°S, 10°W, off the West Coast of Africa, in which Version-6 depicts extensive stratus cloud cover at 1:30 AM, while Version-5 shows very little cloud cover at all. The results shown in Fig. 12a are suggestive of this result because many more cases with p_c greater than 700 mb exist in Version-6 as compared to Version-5. Another noteworthy improvement in Version-6 clouds compared to Version-5 is that

the spatial distribution of clouds in Version-5 has many missing grid points in which no successful cloud retrieval could be performed. There are very few missing grid points (other than orbit gaps) found in Version-6. The percent of grid boxes in which data exist, indicated beneath each figure, shows that Version-6 has retrieved cloud parameter values in roughly 6% more of the grid boxes than does Version-5, both at 1:30 AM and 1:30 PM.

AIRS Outgoing Longwave Radiation (OLR) is computed for each AIRS FOV in which a successful cloud parameter retrieval is performed. QC flags used for OLR are identical to those used for cloud parameters. OLR is computed via a radiative transfer calculation which generates the total longwave flux to space expected for the final retrieved state vector X , including the retrieved cloud parameters⁹. Version-6 OLR has been shown to be of much better quality, and give better agreement with CERES, than does Version-5 OLR⁹. Part of this improvement is a result of the improved accuracy of Version-6 retrieved products as compared to Version-5. In addition, Version-6 uses an improved OLR radiative transfer parameterization¹⁰ compared to what was used in Version-5.

5 Quality Controlled Values of Clear Column Radiances

The clear column radiance for channel ν , τ_{ν} is a derived quantity and, like other Version-6 derived quantities, has case-by-case, channel-by-channel, error estimates $\sigma_{\tau_{\nu}}$, generated in a manner which is described in Appendix A. Version-6 and Version-6 AO use thresholds of $\sigma_{\tau_{\nu}}$ to generate case-by-case, channel-by-channel, QC flags for τ_{ν} , which were not a feature of Version-5. Figure 14 shows statistics over the spectral interval $650 \text{ cm}^{-1} - 760 \text{ cm}^{-1}$, related to QC'd values of τ_{ν} for all oceanic cases within the latitude band $50^{\circ}\text{N} - 50^{\circ}\text{S}$ generated using the 9 day ensemble of retrievals. The top panel of Fig. 14 shows the percent of all cases, as a

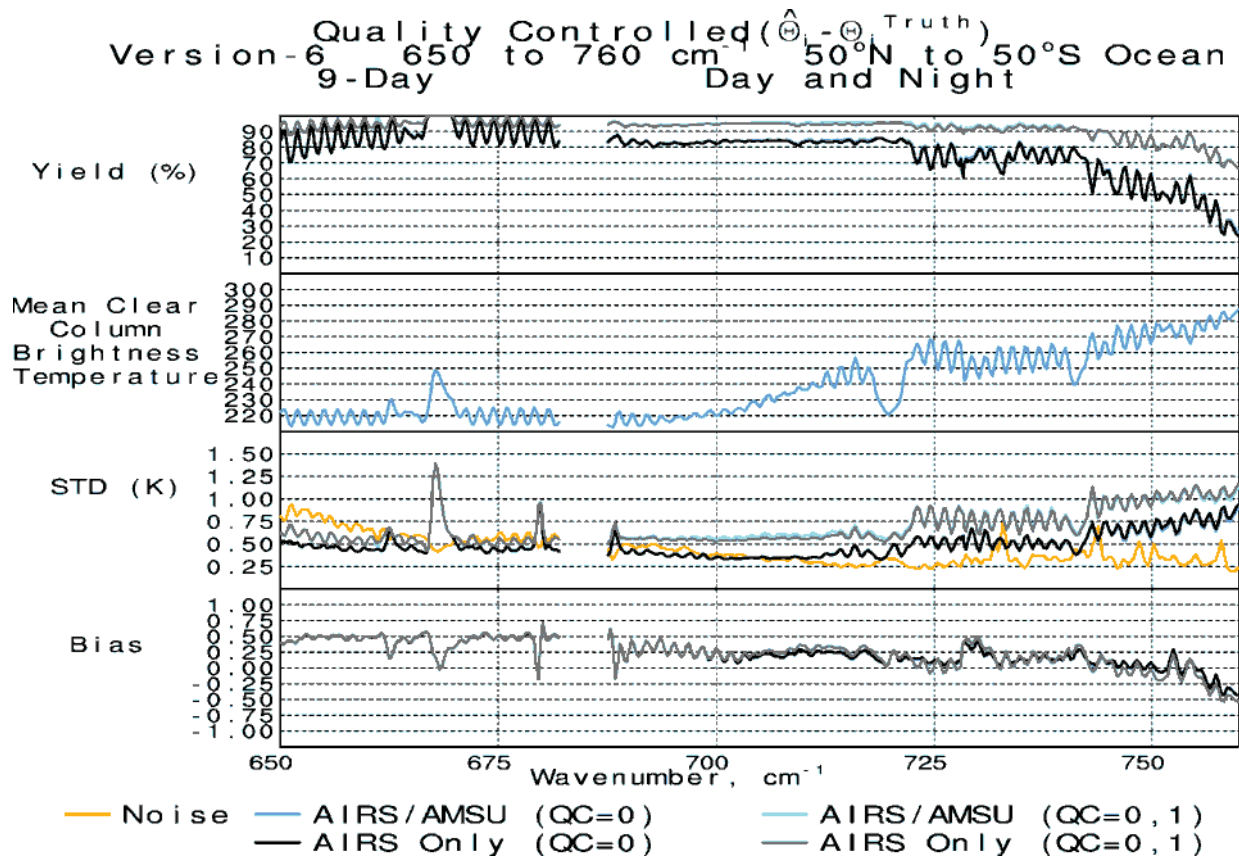


Fig. 14

function of frequency, passing loose Climate thresholds (QC=0,1), and tight DA thresholds (QC=0), in light and dark colors respectively. Results are shown in shades of blue for Version-6 and in shades of black for Version 6-AO. Percent yields are greater for cases passing the Climate QC test as compared to the DA QC test, as expected, but it is important to note that there are no appreciable yield differences between Version-6 and Version-6 AO QC'd values of with regard to either test.

The second panel of Fig. 14 shows the mean values of over all cases with QC = (0,1), where is the clear column brightness temperature given by the blackbody temperature corresponding to . is indicative of the temperature of the portion of the atmosphere to which the channel is most sensitive. Channels with less than 720 cm^{-1} are sensitive primarily to

stratospheric temperatures, and among such channels, those with larger values of ν sound higher in the stratosphere. The reverse is true for channels with ν greater than 720 cm^{-1} , which are sensitive primarily to tropospheric temperatures, and in which higher values of ν indicate increased sensitivity to temperatures in the lower portions of the troposphere, and eventually to the surface skin temperature. Yields of QC'd values of ν generally decrease with increased channel sensitivity to lower tropospheric, and eventually, to surface skin temperatures. Figure 14 shows that yields of accepted values of ν using DA QC thresholds are 50% or higher for channels up to 750 cm^{-1} , which have considerable sensitivity to surface temperature. Yields are higher for those channels sensing higher in the atmosphere, in which observed radiances are less sensitive to cloud cover and ν are therefore less sensitive to cloud clearing errors.

The third panel of Fig. 14 shows the standard deviations (STD) of QC'd values of ν referred to as σ_{ν} , and also shows in yellow, the mean values of the equivalent brightness temperature channel noise σ_{ν} , given by the single FOV channel radiance noise σ_{ν} evaluated at ν . Values of σ_{ν} are computed on a case-by-case basis using the collocated ECMWF state in conjunction with the AIRS Radiative Transfer Algorithm. Errors in both the state and in the AIRS OLR forward calculation will each contribute to errors in σ_{ν} .

. Over land, the surface parameters T_s and ε_v used for truth both contain considerable uncertainty, and they contain some uncertainty over ocean as well. The results shown in Fig. 14 are for ocean cases only, because we do not have accurate estimates of ν over land for channels sensitive to the surface.

Errors in ν arise from two sources: instrumental noise and cloud clearing errors. The channel i clear column radiance ν is obtained as a linear combination of the observed radiances $R_{i,\ell}$ for that channel in each of the $\ell=9$ FOV's used to generate the retrieval. For channels thought to be

unaffected by clouds, the 9 observations are averaged together and this averaging results in a multiplicative channel noise reduction factor of 1/3 for channels that do not “see clouds”. Consequently, the STD’s of $\bar{R}_{i,\ell}$ for stratospheric sounding channels are actually lower than the instrumental noise, especially using DA QC as shown in the darker colors. In general, the taking of a linear combination of $R_{i,\ell}$ to obtain $\bar{R}_{i,\ell}$ amplifies the effect of channel noise on $\bar{R}_{i,\ell}$, especially in the case of more difficult cloud cases¹. Therefore, even if the coefficients used to determine $\bar{R}_{i,\ell}$ from $R_{i,\ell}$ were perfect, the STD of $\bar{R}_{i,\ell}$ would exceed σ_{channel} for channels sensitive to clouds in the FOR. The largest potential source of errors in $\bar{R}_{i,\ell}$ results from errors in the cloud clearing coefficients used to derive $\bar{R}_{i,\ell}$. For both these reasons, the STD of $\bar{R}_{i,\ell}$ increases as frequencies become more sensitive to lower tropospheric and surface temperatures and whose radiances are more greatly affected by clouds. Part of the errors shown at higher ν is an artifact resulting from the effect of the uncertainty in ocean surface skin temperature and ocean spectral emissivity on the values of $\bar{R}_{i,\ell}$ in channels sensitive to the surface. In any event, the STDs of $\bar{R}_{i,\ell}$ using DA QC are not appreciably larger than channel noise up to about 740 cm^{-1} .

The fourth panel of Fig. 14 shows biases of $\bar{R}_{i,\ell}$. Biases of $\bar{R}_{i,\ell}$ for all four ensembles of cases are similar to each other. The small biases outside the higher frequency window region are more likely a result of biases in $R_{i,\ell}$ rather than in $\bar{R}_{i,\ell}$, as well as a result of systematic errors in the RTA. The negative bias of $\bar{R}_{i,\ell}$ in channels more sensitive to the surface may be real and be the result of insufficient cloud clearing when very low clouds are present.

The most important potential application of using QC’d values of $\bar{R}_{i,\ell}$ is with regard to Data Assimilation. ECMWF and NCEP assimilate observed AIRS radiances operationally. In particular, ECMWF and NCEP assimilate AIRS radiances primarily in the spectral interval 650 cm^{-1} - 740 cm^{-1} . These channels are assimilated on a case-by-case, channel-by-channel basis,

using radiances only in those channels whose observed radiances are thought to be unaffected by clouds. In principle, operational centers could assimilate QC'd values of μ in an analogous way given appropriate error estimates and QC procedures. The spatial coverage of QC'd μ is significantly greater than that of radiances unaffected by clouds, especially for tropospheric sounding channels. Figure 14 shows that values of μ with QC=0 over ocean for the most part have yields of 70% or better at frequencies less than 740 cm^{-1} . Moreover, the STD of the errors in μ with QC=0 are on the order of the channel noise at these frequencies. For those cases in which the errors in μ are greater than the channel noise, their individual errors are characterized very well by σ and this can be taken into account by the DA procedure.

6 AIRS Version-6 Data Availability

All AIRS Version-6 and Version-6 AO level-2 and level-3 products can be obtained at the Goddard DISC <http://disc.sci.gsfc.nasa.gov/AIRS/data-holdings>. Spot-by-spot level-2 products are available on an AIRS FOV basis, and gridded level-3 products are presented on a $1^\circ \times 1^\circ$ latitude-longitude grid, gridded separately for 1:30 AM orbits and 1:30 PM orbits, on a daily, eight day mean, and monthly mean basis. More details are given in Olsen et al¹¹ and Manning et al¹².

Appendix A: Error Estimates and QC Procedures

Introduction

Each retrieved quantity X in Version-5 and Version-6 has an associated error estimate σ . A major advancement in Version-5 was the development of methodology to generate accurate empirical error estimates for a number of geophysical parameters, and to use thresholds of these

error estimates for quality control. Analogous procedures are also used in Version-6, with some improvements in the details. Version-4² used threshold values of 12 internal tests for the purpose of generating QC flags for different geophysical parameters. Version-5³ used the case-by-case values of these 12 internal tests as well as values of four additional tests, as predictors to generate case-by-case error estimates for select geophysical parameters X . Version-6 uses methodology to generate empirical error estimates δX which is analogous to that used in Version-5, with some modifications resulting from changes in the steps used in the Version-6 retrieval system compared to those used in Version-5.

A.1 Generation of the empirical error estimates and

Version-5 used case-by-case values of 16 internal tests as predictors in the generation of the empirical error estimates, , and for case . Appendix B of Susskind et al.³ gives a description of these 16 tests. The symbols used for these tests, including their superscripts, have Version-5 heritage, and we maintain the use of the same symbols in the description of the tests used in Version-6. Some of these tests involve procedures used to generate the “start up state” X^l used as the initial guess for the physical retrieval sequence of steps. X^l is also used in the generation of the first pass clear column radiances which are the input to the first phase of the physical retrieval process. The sequences of steps used to generate X^l in Version-5 and Version-6 differ from one another. For this reason, the relevant tests used in each retrieval system are analogous to each other, but refer to results obtained using different states.

In Version-5 and Version-6, and are both computed according to

(A1)

where σ_m is the error estimate of retrieved geophysical parameter for case m , σ_p is the value of the p^{th} test for case m , M is a matrix, and N is the number of predictors used to determine the error estimates. Error estimates are by definition all positive. Values of the predictors are also all positive and in general, larger values of σ_m are indicative that a poorer retrieval will be obtained for case m .

The meanings and significance of the 16 predictors used in equation A1 in Version-5 were as follows: 1) σ_{cl} is the final retrieved effective cloud fraction (%); 2) W_{liq} is cloud liquid water (g/cm^2) retrieved as part of an AMSU Only retrieval step; 3) ΔT_{MW} represents the difference between the retrieved lower-tropospheric temperature obtained in a “test MW only” retrieval step and that determined in the final physical retrieval step (K); 4) σ_{CC} is the channel-noise-amplification factor obtained in the cloud clearing step which generates σ_{CC} (unitless); 5) σ_{CC}^{eff} is the effective channel noise amplification factor obtained in the subsequent and final cloud clearing step that generates σ_{CC}^{eff} , which are the values of clear-column radiances used in the second pass physical retrieval sequence of steps (unitless); 6) σ_{CC}^{fit} represents the quality of the cloud clearing fit obtained in a start-up cloud clearing step used to generate σ_{CC}^{fit} (unitless); 7) R_{temp} represents the degree to which the final physical temperature profile retrieval step has converged (unitless); 8) R_{surf} represents the degree to which the final physical surface-parameter retrieval step has converged (unitless); 9) σ_{CC}^{res} represents the effective channel noise amplification factor resulting from the cloud clearing step used to generate σ_{CC}^{res} (unitless); 10) σ_{CC}^{res} represents the agreement between the observed AMSU channel 5 brightness temperature and that computed from the solution obtained in the final physical retrieval step (K); 11) ΔT_s (2) represents the difference between the final retrieved value of T_s and the value of T_s contained in

the start-up state X^l (K), which is generated by a clear regression step; 12) RS represents the principal component reconstruction score of the observed AIRS radiances that is obtained as part of the start-up clear regression step (unitless); 13) $\Delta T_s^{(1)}$ represents the difference between the final retrieved value of T_s and that contained in X^0 (K), which is generated by a cloudy regression step; 14) ΔT_{low} represents the difference between the retrieved lower tropospheric temperature of the final state and that contained in X^0 ; 15) R_{wat} represents the degree of convergence of the physical retrieval water vapor retrieval step (unitless); and 16) R_{cl} represents the degree of convergence of the cloud clearing step used to generate X^l (unitless).

In Version-6, neither a cloudy regression step nor a clear regression step is used as part of the start-up procedure. These two steps are replaced in Version-6 by a single Neural-Net start-up step which generates X^{NN} . The state X^l used as the initial guess to the physical retrieval process is generated in Version-6 by using X^{NN} as input to an AMSU Only retrieval step, which modifies X^{NN} so as to give X^l . In Version-6 AO, the AMSU Only retrieval step is not performed and X^l is given by X^{NN} . The error estimate predictors used in Version-6 are basically the same as those used in Version-5 with three exceptions: 1) Predictor 10 used in Version-5 was the difference between the observed brightness temperature in AMSU-A channel 5 and the brightness temperature for that channel computed using the final retrieved state. AMSU-A channel 5 has degraded significantly and is no longer used in any way in the Version-6 retrieval process. An analogous predictor is now used in Version-6 involving AMSU-A channel 6. This changes the data used for one predictor in Version-6. 2) Predictor 12 used in Version-5 related to how well the NOAA clear regression step performed. The NOAA clear regression step is not performed in Version-6, and predictor 12 is not computed and therefore not used in the generation of error estimates in Version-6. This eliminates one predictor used in Version-6 from those used in

Version-5. 3) Predictors 11 and 13 used in Version-5 related to the differences between the retrieved surface skin temperature T_s and the skin temperatures obtained using each of the clear and cloudy regression steps in Version-5. In the Neural-Net start up system, there is only one value of surface skin temperature used in the start-up procedure. Therefore, only a single test of this type, given by the difference between the retrieved value of T_s and the value of T_s found in X^{NN} , is used as a predictor in the generation of error estimates in Version-6. This eliminates one additional predictor in Version-6 compared to what was used in Version-5. Consequently, Version-6 uses only 14 error estimate predictors in Equation A1. Version-6 predictors 2, 3, and 10 all involve use of AMSU-A observations in one manner or another and are not used in the Version-6 AO retrieval system. Consequently Version-6 AO uses 11 predictors in the generation of δX . Finally, it has been determined that while the Neural-Net values of T_s are very accurate over ocean, they are not sufficiently accurate over land or frozen ocean for use as an error estimate predictor. Consequently, over land and frozen ocean, the predictor involving the difference between retrieved and Neural-Net surface skin temperature is not used in Version-6. Therefore, over land or sea ice, only 13 predictors are used in Version-6 and 10 predictors are used in Version-6 AO.

The coefficients of M^X are determined in essentially the same manner in Version-6 as was done in Version-5. In Version-6, we generate six distinct matrices M^X for separate use under daytime or nighttime conditions, as well as for separate use over 1) non-frozen ocean; 2) non-frozen land; and 3) frozen (ice or snow) cases. In Version-5, only four such matrices were used, in which a single pair of matrices (day and night) was used to be representative of all cases in categories 2 and 3, and a separate pair of matrices was used over non-frozen ocean. The coefficients of the matrix M^X for an ensemble of cases can be determined in a straightforward

manner if one is given the true values of X , X^{truth} . M^X is determined by finding the coefficients that minimize the RMS difference of $\sum_{i=1}^N (M^X X^{truth} - X)^2$, where $\sum_{i=1}^N (M^X X - X)^2$, when M^X is used in Equation A1 to generate X . The N coefficients of M^X are determined separately for each parameter X . In order to generate the Version-6 coefficients for each of the six different matrices M , we used appropriate spatial subsets of X and X^{truth} , generated using all Version-6 retrievals that were performed on September 29, 2004 and February 24, 2007, along with the collocated ECMWF 3-hour forecast values of X as X^{truth} . The coefficients of the six sets of matrices M were determined separately for Version-6 and Version-6 AO based on observations on these two days, and are then used for all time periods.

In both Version-5 and Version-6, the error estimate for total precipitable water FE is computed in an analogous manner to that used to compute σ and σ_{AO} , but FE is computed in terms of the fractional error estimate

$$FE = \frac{\sum_{i=1}^N (M^W X - X)^2}{\sum_{i=1}^N X^2} \quad (A2)$$

where FE is the fractional error in total precipitable water. The predictors used in Equation A2 are identical to those used in Equation A1. The error estimate for FE is obtained according to Equation A2. The value of total precipitable water X used in Equation A2 is not derived directly in the physical retrieval, but is computed as the vertical integral

The coefficients of M^W are determined in an otherwise analogous way to those of M^X , but by minimizing the RMS difference of $\sum_{i=1}^N (M^W X - X)^2$ when M^W is used in Equation A2.

A 1.1 Non-frozen ocean surface skin temperature Quality Control

Version-5, Version-6, and Version-6 AO all use the non-frozen ocean skin temperature error estimate directly for Quality Control, with separate thresholds and used to indicate best quality retrievals (QC=0) and good quality retrievals (QC=1) respectively. Values of these thresholds are shown in Table A1 for Version-5, Version-6, and Version-6 AO. As in Version-5, in order to achieve a substantial yield of cases with QC=0 or QC=1 poleward of 40°S (lat ≤ -40°), a fixed threshold value was used for latitudes north of 40°S (lat ≥ -40°) and a larger value of was used for latitudes southward of 60° (lat ≤ -60°). The value of used at intermediate latitudes is interpolated linearly in latitude between the two specified values

Table A1 Ocean T_s thresholds (K)

		lat ≥ -40°	lat ≤ -60°
Version-5	0.8	1.0	1.75
Version-6	1.1	1.4	2.0
Version-6 AO	1.2	1.4	2.0

of , both of which are shown in Table A1. Cases in which is less than are considered to have the highest accuracy and are flagged with QC=0. Cases with lying between and are flagged as having good accuracy, with QC=1. Cases with either QC=0 or QC=1 are those used in the generation of the T_s level-3 product over ocean. Cases with are flagged as having poor quality with QC=2. QC flags defined in this manner are what was used in generation of the results shown in Figure 2 of the main text.

Over land or frozen ocean, a different procedure is used for the QC for surface skin temperature and surface spectral emissivities. The reason for this is that ECMWF does not

provide an accurate value of land “truth” to be used in the generation of error estimates, so the error estimates of land surface skin temperature, which are generated analogously to those of ocean surface skin temperature, are less accurate and are not used directly for Quality Control. Surface skin parameter QC flags over land and frozen ocean are generated in the same manner as that used for W_{TOT} as discussed in the main text.

A.1.2 Temperature profile and water vapor profile Quality Control

The methodologies used in Version-5 and Version-6 for the generation of temperature profile QC flags are analogous, but not identical, to each other. As with surface skin temperature, case-by-case level-by-level error estimates for temperature profiles are obtained using Equation A1. These error estimates are subsequently used to determine case-by-case characteristic pressures and down to which the profile is considered to be of highest quality, acceptable for use in data assimilation, or of sufficiently good quality to be used in the generation of level-3 products for climate studies. In Version-5, all IR/MW profiles passing the Stratospheric Temperature Test^{2,3} were assigned to have highest quality (QC=0) down to at least 70 mb. The characteristic pressure p_{best} was defined in Version-5 as the highest pressure (somewhere between 70 mb and p_{surf}) at which the error estimate in each of the next three highest pressure levels is not greater than a pressure dependent error estimate threshold. Temperatures down to p_{best} were assigned the QC flag QC=0.

Pressure dependent thresholds are computed analogously in both Version-5 and Version-6, based on a set of three threshold parameters. These three parameters represent error thresholds defined separately at $p =$, at $p = p_{surf}/2$, and at $p = p_{surf}$, where in Version-5, is 70 mb and in Version-6, is 30 mb. The thresholds

used for QC purposes at intermediate pressures are linearly interpolated in $\log p$ between the appropriate specified values . It was found to be advantageous in Version-5 to have separate temperature profile error thresholds for non-frozen ocean on the one hand, and for land and ice on the other. Version-5 used different sets of thresholds , called Standard Thresholds, for each of these two geographical domains to generate p_{best} as described above. Table A2 shows the values of the Version-5 Standard Thresholds, , used to generate the values of p_{best} consistent with the QC flag QC=0 used in the official Version-5 data set. Table A2 also includes values of the Version-5 Tight Thresholds discussed previously, which were not used for QC flags in the official Version-5 data set. The Version-5 thresholds used over land and ice and snow domains were identical to each other.

As discussed in Section 4.2, over land and frozen ocean, it was found that if one included only those Version-5 cases down to p_{best} , as defined by the Standard Thresholds, in the generation of level-3 products, these level-3 products would have very poor spatial coverage in the lower troposphere over these spatial domains. For this reason, an ad-hoc method was used in Version-5 to define a second characteristic pressure p_{good} which was used to assign the QC flag

Table A2 Temperature Profile Thresholds

	<u>Non-Frozen Ocean</u>			<u>Land</u>			<u>Ice and Snow</u>		
V.5 Standard	1.75	1.25	2.25	2.25	2.0	2.0	2.25	2.0	2.0
V.5 Tight	1.75	0.75	2.0	1.75	0.75	1.75	1.75	0.75	1.75
V.6 DA	3.0	0.75	1.0	3.0	0.75	1.0	3.0	0.75	1.25
V.6 CLIM	3.0	3.0	3.0	3.0	2.0	2.0	3.0	2.5	2.5
V.6 AO DA	3.0	0.8	1.0	3.0	0.85	1.0	3.0	0.85	1.25
V.6 AO CLIM	3.25	3.25	3.25	3.0	2.0	2.0	3.0	2.5	2.5

QC=1 to some additional values of $T(p)$ beneath 300 mb over land and frozen ocean. These additional cases were included in the generation of Version-5 level-3 products, which utilized all cases with QC=0 or QC=1. Temperatures beneath p_{good} were assigned the flag QC=2.

In Version-6 and Version-6 AO, all cases in which the retrieval system converged (about 99% of the cases), are assigned to have highest quality (QC=0) down to at least 30 mb. The characteristic pressures p_{best} and p_{good} , and consequent QC flags of 0, 1, and 2, are defined analogously to what was done in Version-5, with the exception that in Version-6, p_{best} and p_{good} are defined as the lowest pressure for which $\delta q(p)$ exceeded δq_{thr} for N consecutive levels, where N=8 at pressures which are less than 300 mb and N=3 at pressures which are greater than 300 mb, while in Version-5, N=3 at all pressures. In addition, unlike in Version-5, p_{good} in Version-6 is defined using separate sets of δq_{thr} thresholds, referred to as Climate Thresholds, as opposed to those used to define p_{best} , which are referred to as Data Assimilation Thresholds. In Version-6 and Version-6 AO, over land as well as over frozen ocean, if p_{good} as defined above was at most six levels above the surface, corresponding to roughly 1.5 km above the surface, p_{good} was set equal to p_{surf} . Finally, unlike Version-5, Version-6 has separate sets of δq_{thr} thresholds used for cases over ice and snow, which differ slightly from those used over land or non-frozen ocean. Table A2 includes the values of δq_{thr} used in both Version-6 and Version-6 AO.

Error estimates $\delta q(p)$ are generated in a different manner from that used for δT_s , $\delta T(p)$, and W_{TOT} , as described in the next section. The error estimates $\delta q(p)$ are written out as part of the Version-6 data set, but are not used in the generation of QC flags for $q(p)$. In Version-6, the QC flags for $q(p)$ for case m are set to be identical to those of $T(p)$ for that case.

A.2 Generation of empirical error estimates $\delta q(p)$

Error estimates for $q(p)$, $\delta q(p)$, are generated empirically in Version-6 in a manner analogous to what was done in Version-5, according to Equation A3

$$\text{-----} \tag{A3}$$

where are a subset of 7 of the error estimates for derived for that case. The 7 error estimate predictors used in Equation A3 are: 1) $\delta T(150 \text{ mb})^m$; 2) $\delta T(260 \text{ mb})^m$; 3) $\delta T(500 \text{ mb})^m$; 4) $\delta T(750 \text{ mb})^m$; 5) $\delta T(850 \text{ mb})^m$; 6) $\delta T(985 \text{ mb})^m$; and 7) δW_{tot}^m . The coefficients are generated in an analogous fashion to those in Equation A2 using ECMWF values of $q(p)$ as truth and minimizing the RMS fractional errors where is given by

. In Version-6, as with $T(p)$, T_s , and W_{TOT} , coefficients of six distinct matrices, corresponding to daytime and nighttime cases for each of the three geophysical domains described previously, are derived. We use the simplified form of Equation A3 to derive rather than the form of Equation A2, involving more predictors, because we felt that errors in temperature profile and total precipitable water for a given case should be adequate predictors of errors in the water vapor profile. The error estimates which are written out for a given case are computed according to =

A.3 Error estimates and Quality Control for Clear Column Radiances

The general methodology used to analyze AIRS/AMSU observations in both Version-5 and Version-6 is essentially unchanged from that described in Susskind et al.¹. Fundamental to this

approach is the generation of clear column radiances for each AIRS channel i , $\tau_{i,c}$, which are derived products representing the radiance channel i would have seen if the AIRS FOR on which a retrieval is performed were cloud free. $\tau_{i,c}$ is determined for each channel as a linear combination of the observed radiances of that channel in each of the 9 AIRS FOV's contained within the AMSU FOR on which a retrieval is performed according to

$$(A4)$$

where $\bar{\tau}_i$ is the average value of τ_i over the 9 FOV's and \mathbf{a}_i ($i=1,9$) is a derived vector for each FOR obtained as part of the retrieval process. If all values of τ_i used in Equation A4 were perfect, then the error in $\tau_{i,c}$ would be

$$(A5)$$

where ϵ_i is the spatially random noise of channel i and α_i is the channel noise amplification factor, resulting from taking the linear combination of observations in the nine FOV's shown in Equation A4 to obtain $\tau_{i,c}$. It can be shown that the appropriate value of α_i is given by

$$- \quad (A6)$$

Equation A4 shows that $\tau_{i,c}$ if all τ_i are zero. This situation corresponds to a case in which the clear column radiance is obtained by averaging the radiances in all nine FOV's. Equation A6 reduces to $\alpha_i = 1$ when all τ_i are zero. In general, this is not the case and α_i is usually greater than 1, depending on the extent of cloud clearing (extrapolation) performed in the FOR. α_i is in principle channel independent because it arises only from the linear combination of radiances used to construct $\tau_{i,c}$. Some channels are only sensitive to the atmosphere at pressures sufficiently lower than the cloud top pressure (altitudes higher than the cloud top height), and these case dependent channels do not "see" the clouds. The retrieval algorithm determines which channels

do not “see” clouds, and for these channel the retrieval algorithm sets $\tau_{\text{clear}} = 0$ and also sets $\sigma_{\text{clear}} = 0$ for such channels. Equation A6 is used to obtain τ_{clear} for all other channels.

In general, the largest source of noise in τ_{clear} results from errors in the vector \mathbf{y} . In Version-6, as was done in Version-5, σ_{clear} is expressed as the sum of the errors arising from both sources,

$$\sigma_{\text{clear}} = \sqrt{\sigma_{\text{clear,1}}^2 + \sigma_{\text{clear,2}}^2}, \text{ where } \sigma_{\text{clear,1}} \text{ is computed according to} \quad (\text{A7})$$

The seven predictors used in Equation A7 to generate \mathbf{y} are identical to those used in Equation A3 to generate $\mathbf{y}_{\text{clear}}$. In general, the case-by-case clear column radiance error estimate σ_{clear} is computed according to the sum of both sources of noise

$$\sigma_{\text{clear}} = \sqrt{\sigma_{\text{clear,1}}^2 + \sigma_{\text{clear,2}}^2} \quad (\text{A8a})$$

If all 9 values of \mathbf{y} were unaffected by clouds, then for channel i , τ_{clear} would be best approximated by $\tau_{\text{clear,1}}$, and the appropriate value of σ_{clear} for that channel would be given by

$$\sigma_{\text{clear}} = \sigma_{\text{clear,1}} \quad (\text{A8b})$$

In Version-6, Equation A8b is used for τ_{clear} for all FOR's in which $\tau_{\text{clear}} > 0$ for channel i , and Equation A8a is used otherwise.

Clear column radiances and their associated error estimates are written out in radiance units ($\text{W}/\text{m}^2\text{-sr}\text{-cm}^{-1}$). It is more convenient, however, to think in terms of clear column brightness temperatures T_{clear} , and their error estimates $\sigma_{T_{\text{clear}}}$, both given in K. The clear column brightness temperature T_{clear} is the equivalent blackbody temperature of T_{clear} , defined as the temperature for which $\tau_{\text{clear}} = \tau_{\text{clear,1}}$ where $B_{\lambda}(T)$ is the Planck blackbody function. Given τ_{clear} and σ_{clear} , T_{clear} is evaluated according to

$$\tau_{\text{clear}} = \tau_{\text{clear,1}} \quad (\text{A9})$$

As in the generation of other empirical error estimates, Version-6 used six different matrices M^R in Equation A8a, one for each of six different spatial and temporal domains. The coefficients of the six different matrices M^R were determined analogously to those of the other matrices M described previously, such that the coefficients minimize the RMS differences of

where ϵ and σ is computed using Equations A8a and A9.

The true clear column brightness temperature T_{bc} is defined as the value of T that is computed using the AIRS Radiative Transfer Algorithm in conjunction with the truth state X^{truth} .

We used ECMWF 3-hour forecast values for T , q , and τ . The

profile used in the calculations had a spatially homogeneous vertically constant mixing ratio which was set to be 371.79 ppm on January 1, 2002, and increased linearly in time at a rate of 2.026 ppm/yr. The truth values used for $CO(p)$, and $CH_4(p)$ were based on spatially varying monthly mean climatologies. The AIRS Team model was used as truth for surface emissivity over non-frozen ocean. Reasonable globally homogeneous surface emissivity values were used as truth over land. Values of ϵ_{CO_2} , and therefore of ϵ_{CH_4} are most accurate for channels in the 15 μm and 4.3 μm CO_2 bands, especially those channels which are less sensitive to surface emission. For this reason, the best error estimate coefficients and error estimates are generated in the 15 μm and 4.3 μm CO_2 bands for those channels which are not sensitive to surface emission. Error estimate coefficients generated for channels which are very sensitive to water vapor or ozone absorption are less accurate because of limitations in the truth values used for water vapor and ozone profiles. Error estimate coefficients for those channels which are very sensitive to surface emission are also less reliable, but are better over ocean than over land. Finally, clear column radiances at frequencies greater than or equal to 2175 cm^{-1} are affected by incoming solar radiation reflected by the surface back in the direction of the satellite. The relevant surface

bi-directional reflectance term is not modeled well in the computation of . For this reason, daytime values of for frequencies between 2180 cm^{-1} and 2240 cm^{-1} , and between 2380 cm^{-1} and 2660 cm^{-1} , would be of lower accuracy because radiances in these channels are sensitive to reflected solar radiation which is not well modeled in . Therefore, we substituted the values of determined during nighttime conditions for these channels, in place of those that were computed during daytime conditions, in the daytime matrices of .

A.3.1 QC flags for

Different channels are sensitive, by varying amounts, to clouds at different pressures. Therefore, is both channel and case dependent. Even if significant cloud clearing errors exist for some channels in a given case, channels that have little or no sensitivity to the clouds in that case would have very accurate values of . It is for this reason that we assign each channel its own case dependent QC flags indicating whether the cloud-cleared radiance is of sufficient accuracy for use for different purposes. We used the predicted clear column brightness temperature error to assign the QC flags for on a case-by-case basis. In Version-6, is assigned the flag QC=0 if is less than 1.0K, and is assigned the QC flag QC=1 if is between 1.0K and 2.5K. Otherwise, the QC flag is set equal to 2. The flag QC=0 is intended to mark those channels that are thought to be accurate enough for data assimilation purposes, with the goal that the error in should be not much larger than the channel noise . The flag QC=1 is designed to provide better spatial coverage for a given channel for use in process studies, but still eliminate poor values of . Figure 14 of the main text shows acceptance yields and RMS errors of QC'd values of from 650 cm^{-1} – 760 cm^{-1} in which the QC procedures used are as defined in this section.

References

1. J. Susskind, C. D. Barnet, and J. M. Blaisdell, "Retrieval of atmospheric and surface parameters from AIRS/AMSU/HSB data in the presence of clouds," *IEEE Trans. Geoscience and Rem. Sens.*, **41**, 390-409 (2003) [doi: 10.1109/TGRS.2002.808236].
2. J. Susskind, C. D. Barnet, J. M. Blaisdell, L. Iredell, F. Keita, L. Kouvaris, G. Molnar, and M. T. Chahine, "Accuracy of geophysical parameters derived from Atmospheric Infrared Sounder/Advanced Microwave Sounding Unit as a function of fractional cloud cover," *J. of Geophys. Res.*, **111**, D09S17 (2006) [doi:10.1029/2005JD006272].
3. J. Susskind, J. M. Blaisdell, L. Iredell, and F. Keita, "Improved temperature sounding and quality control methodology using AIRS/AMSU data: The AIRS Science Team Version-5 Retrieval Algorithm," *IEEE Trans. on Geoscience and Rem. Sens.*, **44**, 883-907 (2011) [doi: 10.1109/TGRS.2010.2070508].
4. M. T. Chahine, "Remote sensing of cloudy atmospheres. I. The single cloud layer," *J. Atmos. Sci.*, **31**, 233-243 (1974).
5. M. T. Chahine, "Remote sensing of cloudy atmospheres. II. Multiple cloud formations," *J. Atmos. Sci.*, **34**, 744-757 (1977).
6. Z. Tao, W. J. Blackwell, D. H. Staelin, "Error variance estimation for individual geophysical parameter retrievals," *IEEE Trans. Geoscience and Rem. Sens.*, **51**, 1718-1727 (2013) [doi:10.1109/TGRS.2012.2207728].
7. W. J. Blackwell, "Neural network Jacobian analysis for high-resolution profiling of the atmosphere," *EURASIP Journal on Advances in Signal, Processing*, 2012, 2012.71 [doi:10.1186/1687-6180-2012-71].

8. J. Susskind and O. Reale, “Improving forecast skill by assimilation of AIRS temperature soundings,” *Geoscience and Rem. Sens. Symp. (IGARSS)*, 2010 IEEE Intl., July 25-30, 3534-3537 (2010) [doi:10.1109/IGARSS.2010.5652458].
9. J. Susskind, G. Molnar, L. Iredell, and N. G. Loeb, “Interannual Variability of Outgoing Longwave Radiation as observed by AIRS and CERES,” *J. Geophys. Res.*, **117**, D23107 (2012) [doi:10.1029/2012JD017997].
10. M. J. Iacono, J. S. Delamere, E. J. Mlawere, M. W. Shephard, S. A. Clough, and W. D. Collins, “Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models,” *J. Geophys. Res.*, **113**, D13103 (2008) [doi:10.1029/2008JD009944].
11. E. T. Olsen, E. Manning, S. Licata, J. Blaisdell, L. Iredell, J. Susskind, “AIRS/AMSU/HSB Version 6 Data Release User Guide,” Jet Propulsion Laboratory, Pasadena, CA (2013).
12. E. M. Manning, S. Z. Friedman, “AIRS Version 6.0 released files description,” Jet Propulsion Laboratory, Pasadena, CA (2013).

Caption List

Fig. 1 Sample AIRS cloud free brightness temperature spectrum. The channels used in different retrieval steps in Version-5 are indicated by stars of different colors.

Fig. 2 Statistics of QC'd SST differences from ECMWF “truth” for Version-5, Version-6, and Version-6 AO using each DA QC and Climate QC thresholds.

Fig. 3 Nine-day mean difference of Version-6 and Version-5 level-3 SST products from colocated ECMWF truth for ocean grid points between 50°N and 50°S.

Fig. 4 Statistics related to ocean surface emissivity as a function of satellite zenith angle, at both 950 cm^{-1} and 2400 cm^{-1} .

Fig. 5 Difference of 1:30 AM and 1:30 PM nine-day mean land level-3 emissivity products shown at 950 cm^{-1} and 2400 cm^{-1} for each of Version-6 and Version-5.

Fig. 6 Global mean statistics of QC'd Version-5, Version-6, and Version-6 AO temperature profiles, compared to ECMWF truth, using different QC thresholds.

Fig. 7 Comparison of the accuracies of QC'd Version-6 and Version-5 retrieved temperature profiles with those of their initial guesses.

Fig. 8 a) Percent acceptance, using Climate QC, of global Version-5 and Version-6 temperature profiles as a function of retrieved fractional cloud cover at three select pressure levels. b) RMS difference of Version-5 and Version-6 1 km layer mean temperatures from co-located truth in three select layers.

Fig. 9 Global mean yield and spurious layer mean temperature bias trends of QC'd Version-5, Version-6, and Version-6 AO retrievals as a function of pressure.

Fig. 10 Global mean statistics of QC'd Version-5, Version-6, and Version-6 AO water vapor profiles, compared to ECMWF truth, using different QC thresholds.

Fig. 11 Difference of Climate QC'd nine-day mean total precipitable water (cm) from colocated ECMWF for Version-6 and Version-5.

Fig. 12 Global statistics of Version-5, Version-6, and Version-6 AO cloud parameter retrievals as a function of retrieved cloud top pressure: a) Number of retrieved cases for a given cloud top pressure; b) Average retrieved cloud fraction of a function of cloud top pressure.

Fig. 13 Version-6 and Version-5 retrieved cloud fractions and cloud top pressures for 1:30 AM and 1:30 PM orbits on September 29, 2004.

Fig. 14 Statistics for Quality Controlled Version-6 and Version-6 AO cloud cleared brightness temperatures over the spectral interval 650 cm^{-1} to 760 cm^{-1} , using two sets of QC thresholds. Results shown are for all accepted oceanic cases 50°N to 50°S .