

# Improved Temperature Sounding and Quality Control Methodology Using AIRS/AMSU Data: The AIRS Science Team Version 5 Retrieval Algorithm

Joel Susskind, John M. Blaisdell, Lena Iredell, and Fricky Keita

**Abstract**—This paper describes the AIRS Science Team Version 5 retrieval algorithm in terms of its three most significant improvements over the methodology used in the AIRS Science Team Version 4 retrieval algorithm. Improved physics in Version 5 allows for use of AIRS clear column radiances in the entire 4.3  $\mu\text{m}$   $\text{CO}_2$  absorption band in the retrieval of temperature profiles  $T(p)$  during both day and night. Tropospheric sounding 15  $\mu\text{m}$   $\text{CO}_2$  observations are now used primarily in the generation of clear column radiances  $\hat{R}_i$  for all channels. This new approach allows for the generation of more accurate values of  $\hat{R}_i$  and  $T(p)$  under most cloud conditions. Secondly, Version 5 contains a new methodology to provide accurate case-by-case error estimates for retrieved geophysical parameters and for channel-by-channel clear column radiances. Thresholds of these error estimates are used in a new approach for Quality Control. Finally, Version 5 also contains for the first time an approach to provide AIRS soundings in partially cloudy conditions that does not require use of any microwave data. This new AIRS Only sounding methodology, referred to as AIRS Version 5 AO, was developed as a backup to AIRS Version 5 should the AMSU-A instrument fail. Results are shown comparing the relative performance of the AIRS Version 4, Version 5, and Version 5 AO for the single day, January 25, 2003. The Goddard DISC is now generating and distributing products derived using the AIRS Science Team Version 5 retrieval algorithm. This paper also described the Quality Control flags contained in the DISC AIRS/AMSU retrieval products and their intended use for scientific research purposes.

**Index Terms**—AIRS/AMSU, high spectral resolution IR sounders, retrieval methodology, IR sounding in cloudy conditions, cloud cleared radiances, error estimates.

## I. INTRODUCTION

AIRS was launched on EOS Aqua on May 4, 2002, together with AMSU-A and HSB, to form a next generation polar orbiting infrared and microwave atmospheric sounding

Manuscript received. (Write the date on which you submitted your paper for review.)

J. Susskind is with the National Aeronautics and Space Administration Goddard Space Flight Center (NASA GSFC), Greenbelt, MD 20771 USA (phone: 301-286-7210; fax: 301-286-2323; email: [Joel.Susskind-1@nasa.gov](mailto:Joel.Susskind-1@nasa.gov)).

J. Blaisdell is with SAIC at NASA GSFC, Greenbelt, MD 20771 USA (email: [John.Blaisdell@nasa.gov](mailto:John.Blaisdell@nasa.gov)).

L. Iredell is with SAIC at NASA GSFC, Greenbelt, MD 20771 USA (email: [Lena.Iredell@nasa.gov](mailto:Lena.Iredell@nasa.gov)).

F. Keita was with SAIC at NASA GSFC, Greenbelt, MD 20771 USA.

system [1]. The sounding goals of AIRS are to produce 1 km tropospheric layer mean temperatures with an rms error of 1K, and 1 km layer precipitable water with an rms error of 20 percent, in cases with up to 80 percent effective cloud cover. The primary products of AIRS/AMSU-A are twice daily global fields of atmospheric temperature-humidity profiles, ozone profiles, sea/land surface skin temperature, and cloud related parameters including OLR. Also included are the clear column radiances used to generate these products. These are a derived quantity representative of the radiances AIRS would have seen if there were no clouds in the field of view. All products also have error estimates. The products are designed for data assimilation purposes so as to improve numerical weather prediction, as well as for the study of climate and meteorological processes. With regard to data assimilation, one can use either the products themselves or the clear column radiances from which the products were derived.

The theoretical approach used to analyze AIRS/AMSU/HSB data in the presence of clouds in the Version 3 at-launch algorithm, and that used in the Version 4 post-launch algorithm, have been described previously [2,3]. This paper describes the AIRS Science Team Version 5 algorithm, now being used operationally at the Goddard DISC. The AIRS Version 5 retrieval algorithm contains three significant theoretical improvements over Version 4. Improved physics in Version 5 allows for use of AIRS clear column radiances in the entire 4.3  $\mu\text{m}$   $\text{CO}_2$  absorption band in the retrieval of temperature profiles  $T(p)$  during both day and night. Tropospheric sounding 15  $\mu\text{m}$   $\text{CO}_2$  observations are now used primarily in the generation of clear column radiances  $\hat{R}_i$  for all channels. This new approach allows for the generation of more accurate values of  $\hat{R}_i$  and  $T(p)$  under most cloud conditions. Secondly, Version 5 also contains a new methodology to provide accurate case-by-case error estimates for retrieved geophysical parameters and for channel-by-channel clear column radiances. Thresholds of these error estimates are used in a new approach for Quality Control. Finally, Version 5 also contains for the first time an approach to provide AIRS soundings in partially cloudy conditions that does not require use of any microwave data. This new AIRS Only sounding methodology, referred to as AIRS Version 5 AO, was developed as a backup to AIRS Version 5 should the AMSU-A instrument fail. (The HSB instrument, containing the 183 GHz water vapor sounding channels, failed in February 2003. For consistency of results over time, no HSB channels are used in the Version 5 retrieval algorithm as run at the Goddard DISC in any time period.)

## II. OVERVIEW OF THE AIRS VERSION 5 RETRIEVAL ALGORITHM

The AIRS Science Team Version 5 and Version 5 AO retrieval algorithms are very similar to each other and to those described previously [2,3]. The key steps are outlined below: 1) Start with an estimated initial state consistent with the observed radiances; 2) Derive IR cloud cleared radiances,  $\hat{R}_i^0$ , valid for the 3x3 AIRS Fields of View (FOVs) within an AMSU-A Field of Regard (FOR) consistent with the estimated initial state and the observed radiances using 58 AIRS cloud clearing channels; 3) Obtain an AIRS regression guess  $X^{\text{reg}}$  [4] consistent with  $\hat{R}_i^0$  using 1504 AIRS channels; 4) Derive  $\hat{R}_i^1$  consistent with the AIRS radiances and  $X^{\text{reg}}$ ; 5) Derive all surface and atmospheric parameters using  $\hat{R}_i^1$  for 308 AIRS channel radiances and AMSU radiances (Version 5 AO is otherwise identical but does not use the AMSU observations) using the pre-launch physical retrieval methodology which is comprised of a number of sequential steps [2]; 6) Derive an improved set of cloud cleared radiances  $\hat{R}_i^2$  consistent with the observed radiances and the AIRS physically retrieved state parameters; 7) Repeat Step 5 using  $\hat{R}_i^2$  to produce the final retrieval state; 8) Derive cloud parameters and OLR consistent with the solution and observed radiances  $R_i$ ; 9) Apply initial Quality Control, which rejects the final solution if the retrieved cloud fraction is greater than 90% or other relatively coarse tests fail. In the event that a retrieval is rejected, cloud parameters and OLR are determined consistent with the state used for initial cloud clearing, in conjunction with the observed AIRS radiances. Otherwise, cloud parameters and OLR are computed using the final retrieval and observed AIRS radiances, and further Quality Control is subsequently applied to individual geophysical parameters.

The major structural differences between the Version 5 and Version 4 algorithms are related to the new ability to perform cloud clearing using only AIRS observations and the new methodology to determine accurate case-by-case, parameter-by-parameter error estimates. These differences are described in the next two sections. Version 5 and Version 5 AO are otherwise identical except that Version 5 AO does not use AMSU-A radiances in any step, including the generation of error estimates and their use for Quality Control.

## III. BASIC THEORETICAL CONSIDERATIONS FOR CLOUD CLEARING

Using the assumption that adjacent fields of view have otherwise identical geophysical conditions except for cloud cover, Chahine [5] has shown that in the case of  $K-1$  cloud formations, observations in  $K$  fields of view (FOV's) are needed to obtain channel  $i$  cloud cleared radiances  $\hat{R}_i$  according to

$$\hat{R}_i = R_{i,1} + \sum_{k=1}^{K-1} \eta_k (R_{i,1} - R_{i,K+1-k}) \quad (1)$$

where  $R_{i,k}$  is the channel  $i$  radiance observation in field of view  $k$ . A set of  $I$  equations of the general form of Equation 1,

one for each cloud clearing channel  $i$ , are used to determine the vector  $\eta_k$ . The parameters  $\eta_k$  determined in Equation 1, using  $I$  channels, are assumed to characterize the cloud formations and thus should be valid for use in all channels.

In analyzing AIRS/AMSU-A data, a single sounding is produced using all 9 AIRS FOV's falling within a single AMSU-A footprint [2,3]. The AIRS Science Team has found it is advantageous to extrapolate the radiances in the 9 fields of view according to a similar equation of the form

$$\hat{R}_i^n = R_{i,\text{AVG}} + \sum_{k=1}^9 \eta_k^n (R_{i,\text{AVG}} - R_{i,k}) \quad (2)$$

where  $R_{i,\text{AVG}}$  is the average channel  $i$  radiance of all 9 fields of view. The superscript  $n$  has been added to Equation 2 because the generation of  $\hat{R}_i$  is iterative as described in Section II. Optimal values of  $\eta_k^n$  will give true values of  $\hat{R}_i$  up to instrumental noise effects. While there are 9 values of  $\eta_k$  shown in equation 2, only 8 of them are linearly independent.

Most of the  $I=58$  cloud clearing channels used in Equation 1 are channels in the 15  $\mu\text{m}$   $\text{CO}_2$  absorption band with temperature weighting functions covering the atmosphere from the lower stratosphere to the surface. Observations in some window channels are used as well as the determination of  $\eta_k^n$ .

### A. Determination of $\eta$

If, for each channel  $i$ , one substitutes an estimate of  $R_{i,\text{CLR}}^n$  for  $\hat{R}_i^n$  in Equation 2, this gives  $I$  equations for  $K (=9)$  unknowns. The unconstrained weighted least square solution to this multilinear problem is given by

$$\eta_{K \times 1}^n = \left[ \Delta R' N^{-1} \Delta R \right]_{K \times K}^{-1} \Delta R' N^{-1} \Delta R^n_{\text{CLR}} \quad (3)$$

where  $\Delta R$  is an  $I \times K$  matrix with  $\Delta R_{i,k} = R_{i,\text{AVG}} - R_{i,k}$ ,  $\Delta R^n_{\text{CLR}}$  is an  $I \times 1$  matrix given by  $\Delta R^n_{i,\text{CLR}} = R_{i,\text{CLR}}^n - R_{i,\text{AVG}}$ , and  $N$  is an  $I \times I$  channel noise covariance matrix.  $R_{i,\text{CLR}}^n$  is generated by computing expected radiances for cloud clearing channel  $i$  based on the current estimate of the geophysical state  $X^n$  assuming cloud free conditions. As in Susskind et al. [2,3], the solution for  $\eta$  is stabilized by solving for coefficients of up to the first four principal components of  $\left[ \Delta R' N^{-1} \Delta R \right]$ , solving for the coefficient of a principal component only if its eigenvalue is sufficiently large.

The key to the accurate determination of  $\eta$  is obtaining the best estimates of  $\Delta R^n_{i,\text{CLR}}$ , along with an accurate treatment of the noise covariance matrix  $N$ . The values of  $\Delta R^n_{i,\text{CLR}}$  which are used to determine  $\eta$  (and  $\hat{R}_i$ ) are iterative and are computed based on the current best estimate of all relevant surface and atmospheric properties. It is best for the estimated geophysical parameters to be unbiased over

large regions of the atmosphere. In all previous versions of the cloud clearing algorithm, including Version 4, the geophysical state  $X^0$  used to estimate  $R_{i,CLR}^0$  was derived from an AMSU-A retrieval state [6], thus insuring an unbiased temperature and moisture profile over coarse layers in the atmosphere. Subsequent research has shown that a reasonably good regression relationship can be obtained between geophysical parameters and observed (un-cloud cleared) AIRS radiances<sup>1</sup>, and this regression based state can be used to generate an alternative initial state  $X^0$  used for initial cloud clearing. Generation of this state  $X^0$  does not require use of any AMSU observations. This is the approach used to obtain  $X^0$  in the AIRS Only cloud clearing system, Version 5 AO. Version 5 uses analogous methodology to give the initial cloud clearing state, but the cloudy regression also makes use of AMSU observations. The state derived from this cloudy regression,  $X^{reg}$ , is followed in Version 5 by an AMSU only temperature profile retrieval step to produce the state  $X^0$  which is used for initial cloud clearing. Such a step is not performed in Version 5 AO.

The cloudy regression can produce biased initial states, especially if it is not followed by an AMSU-A retrieval step. If the state  $T^0(p)$  used to derive  $R_{i,CLR}^0$  were biased (say too warm), incorrect values of  $\eta_k^0$  would be determined which would result in  $\hat{R}_i^0$  being too large, which in turn would result in the retrieved  $T^1(p)$  being too warm. Chahine [7] has shown that it is optimal to use only longwave (15  $\mu\text{m}$  and 12  $\mu\text{m}$ ) channels for cloud clearing, and shortwave 4.2  $\mu\text{m}$  channels for the determination of temperature profiles beneath the clouds. This is done so as to minimize the bias in retrieved temperature profiles  $X^{n+1}$  resulting from biases in the temperature profile  $X^n$  used to determine  $\eta_k$ . Up to Version 4, most 4.2  $\mu\text{m}$  channels could not be utilized during the day because these channels are affected by non-Local Thermodynamic Equilibrium (non-LTE) which was not accounted for in the Version 4 Radiative Transfer Algorithm (RTA) [3]. Therefore, the optimal combination of channels for cloud clearing and sounding purposes was not used in Version 4, in which most temperature sounding channels were in the 15  $\mu\text{m}$  CO<sub>2</sub> band.

A significant improvement over the AIRS Science Team Version 4 retrieval algorithm is the use of a new RTA which contains improved characterization of atmospheric absorption characteristics, and more significantly, accounts for effects of non-LTE [8]. This for the first time enables the use of all shortwave CO<sub>2</sub> channels in the temperature profile retrieval step of the physical retrieval algorithm, both day and night. This new RTA also contains a better parameterization of the absorption characteristics of the AIRS channels as a function of atmospheric temperature and constituent profile as a function of satellite zenith angle.

<sup>1</sup>AIRS cloudy radiance regression coefficients were provided by Chris Barnett of NOAA. The methodology used to generate these coefficients is analogous to that in [4], but uses observed radiances rather than cloud cleared AIRS radiances.

#### IV. CHANNELS USED IN DIFFERENT STEPS OF VERSION 5

Fig. 1 shows a typical AIRS brightness temperature spectrum and includes the channels used in Version 5 and Versions 5 AO for cloud clearing, and in each of the different steps of the AIRS physical retrieval algorithm. These channels are summarized below.

##### A. Temperature profile retrievals

The location and number of channels used in the temperature profile retrieval step (red stars) are significantly different from those of Version 4 [3]. The major difference is in the incorporation of non-LTE CO<sub>2</sub> temperature sounding channels in the spectral region 2358  $\text{cm}^{-1}$ –2386  $\text{cm}^{-1}$  in the temperature profile retrieval step, which now uses 49 channels between 2197  $\text{cm}^{-1}$  and 2395  $\text{cm}^{-1}$  that are sensitive to both stratospheric and tropospheric temperatures, as well as 42 stratospheric sounding channels between 664  $\text{cm}^{-1}$  and 712  $\text{cm}^{-1}$ . Tropospheric sounding longwave CO<sub>2</sub> channels are now used only for cloud clearing (yellow stars), and are no longer used in the temperature profile retrieval step. These same channels are also the ones used in the cloud parameter retrieval step [2]. Version 5 also includes appropriate AMSU-A channels in the temperature profile retrieval step, as does Version 4 [3], while Version 5 AO does not.

##### B. Surface parameter retrievals

As in Version 4, 25 channels are used in the surface parameter step (blue stars), 15 of which are between 759  $\text{cm}^{-1}$  and 1228  $\text{cm}^{-1}$ , and the remainder are between 2450  $\text{cm}^{-1}$  and 2659  $\text{cm}^{-1}$ . These channels are used to determine simultaneously surface skin temperature  $T_s$ , surface spectral emissivity  $\epsilon_v$ , and short-wave surface bi-directional reflectance  $\rho_v$ .

##### Spectral emissivity perturbation functions

As in Version 4, given an initial surface spectral emissivity guess  $\epsilon_v^0$ , the final surface spectral emissivity  $\epsilon_v$  is expressed as

$$\epsilon_v = \epsilon_v^0 + \sum_1^{N_F} A_i F_i(v) \quad (4)$$

and the final surface bi-directional reflectance  $\rho(v)$  is expressed as

$$\rho_v = \rho_v^0 + \sum_{i=1}^{N_G} B_i G_i(v) = \frac{(1 - \epsilon_v^0)}{\pi} + \sum_{i=1}^{N_G} B_i G_i(v) . \quad (5)$$

$N_F$  and  $N_G$  in Equations 4 and 5 are the numbers of spectral emissivity and spectral surface bi-directional reflectance functions being solved for in the physical retrieval step. Surface bi-directional reflectance is solved for only during the day. Therefore, including the surface skin temperature, which is also solved for in the surface parameter retrieval step, a total of  $1 + N_F + N_G$  unknown coefficients are solved for during the day and  $1 + N_F$  unknown coefficients are solved for at night. The emissivity perturbation functions  $F$  and  $G$  are triangles

linear in frequency. In the case of a single function in a spectral region, the correction to the initial guess is constant in frequency. Otherwise, the correction is piecewise linear. The methodology used to determine the unknown coefficients in this step, and in all retrieval steps, is identical to that described previously [2].

In Version 4,  $N_F = 2$  and there is 1 function covering the longwave portion of the spectrum and 1 function covering the shortwave portion of the spectrum, with unknown coefficients  $A_1(A_{lw})$  and  $A_2(A_{sw})$ . In Version 5,  $N_F = 4$ , with coefficients of 3 perturbation functions being solved for as well as of 1 shortwave emissivity perturbation function. In both Version 4 and Version 5,  $N_G = 1$ .

Over non-frozen ocean,  $\epsilon_v^o$  is set equal to the values found in the AIRS Science Team ocean emissivity model. That model is re-interpolated from calculations by Paul VanDelst [9] using algorithms by Wu referenced in Wu and Smith [10].

### C. Constituent profile retrievals

As with earlier versions, constituent profile retrievals are performed in separate steps, each having their own set of channels and functions [2,3]. In Version 4, water vapor profile retrievals  $q(p)$ , and ozone profile retrievals  $O_3(p)$  were performed after the first pass temperature retrieval  $T(p)$  and carbon monoxide profile retrievals  $CO(p)$  were performed after the final  $T(p)$  retrieval [3]. Version 5 constituent profile retrievals are performed in an analogous manner to those of Version 4, with small modifications made to the channels and functions used in the retrieval process. Version 5 also contains a new retrieval step to determine methane profile  $CH_4(p)$  performed after the final  $T(p)$  retrieval step in an analogous manner to all other constituent profile retrieval steps. Fig. 1 shows in different colors the Version 5 channels used in each of these retrieval steps. The  $q(p)$  retrieval (pink stars) uses 33 channels in the spectral ranges  $1377\text{ cm}^{-1}$  to  $1605\text{ cm}^{-1}$  and  $2608\text{ cm}^{-1}$  to  $2656\text{ cm}^{-1}$ ; the  $O_3(p)$  retrieval (green stars) uses 41 channels between  $997\text{ cm}^{-1}$  and  $1069\text{ cm}^{-1}$ ; the  $CO(p)$  retrieval uses 36 channels between  $2181\text{ cm}^{-1}$  and  $2221\text{ cm}^{-1}$ ; and the  $CH_4(p)$  retrieval (brown stars) uses 71 channels between  $1230\text{ cm}^{-1}$  and  $1356\text{ cm}^{-1}$ .

In Version 4, the first guesses used in each  $q(p)$  retrieval step and the  $O_3(p)$  retrieval steps came from the regression step, while climatology was used in the  $CO(p)$  retrieval step. In Version 5, the output of the regression step is still used as the first guess for the  $q(p)$  retrieval, while climatology is used as the first guess for each of the  $O_3(p)$ ,  $CO(p)$ , and  $CH_4(p)$  retrieval steps.

### D. Version 5 AIRS tuning coefficients

As with Version 4, there are still biases between observed cloud free brightness temperatures for channel  $i$  and those computed with the new RTA using the “true” surface and atmospheric state. These biases, called tuning coefficients, are added to the computed brightness temperatures used in each of the steps in Version 5. These tuning coefficients are shown for

all channels used in the retrieval process in Fig. 1. They are considerably smaller than those used in Version 4 [3], and are generally less than 0.3K. No tuning is applied in the  $CH_4(p)$  retrieval step. Tuning coefficients for channels used in this step are shown as zero in Fig.1.

## V. ERROR ESTIMATES AND QUALITY CONTROL

### A. Approach used in Version 4 for Quality Control

Coupled AIRS/AMSU-A (or AIRS only) retrievals in the presence of broken cloud cover are usually highly accurate. Under some conditions, such as complete overcast, combined AIRS/AMSU-A retrievals cannot be performed at all. In cases of complex clouds or terrain, retrievals are of poorer quality. In the pre-launch version of the AIRS/AMSU-A retrieval algorithm, Quality Control was applied uniformly to the entire profile. If any geophysical parameter was considered to be of poor quality, the whole set of retrieval geophysical parameters was rejected and clouds were derived using the MW retrieval state obtained in Step (1) above. This “one size fits all” approach led to significant compromises between desired spatial coverage of accepted retrievals and desired accuracy. In Version 4 [3], the combined IR/MW retrieval parameters were retained, and used to derive cloud parameters, as long as it was felt that the combined IR/MW retrieval (Step 7) was at least as accurate as the MW only retrieval (Step 1). This was considered to be true if the retrieved cloud fraction derived using the IR/MW state was less than or equal to 90% and the initial cloud clearing step was stable. If this test was passed (referred to in Version 4 as the Stratospheric Temperature Test), the temperature profile above 200 mb was considered acceptable. Constituent profiles ( $H_2O$ ,  $O_3$ ,  $CO$ , and  $CH_4$ ) were accepted if the Stratospheric Temperature Test was passed and additional slightly more stringent cloud clearing stability tests were also passed. The next level of test in Version 4 was applied to the temperature profile beneath 200 mb and above 3 km (the Mid Tropospheric Temperature Test). Finally, a more stringent test was applied to accept temperature profiles in the lowest 3 km of the atmosphere (Lower Tropospheric Temperature Test). Lower tropospheric temperatures are the most difficult to determine accurately, because of effects of low clouds on the radiances, as well as uncertainty and small scale variability in surface skin temperature and emissivity. Both concerns create greater problems over land than ocean. As a result of this, the Lower Tropospheric Temperature Test rejected lower tropospheric temperatures considerably more often over land than over ocean in Version 4. Finally, Version 4 had two additional tests applied over ocean for SST Quality Control, called the Standard SST test and the Tight SST test, respectively.

### B. Version 5 error estimates and Quality Control for retrieved temperature profiles, surface skin temperatures, and total precipitable water

The methodology used in Version 4 for quality control represented a significant improvement over the previously used “one size fits all” quality control methodologies which classified an entire sounding as either acceptable or

unacceptable. Nevertheless, the Version 4 quality control methodology was totally ad-hoc and was based on whether soundings passed different sets of threshold criteria. Version 5 uses an improved Quality Control methodology based on thresholds of the error estimates of derived geophysical parameters.

The Version 4 quality control used thresholds for values of 12 different parameters  $Y_k$  ( $k=1,12$ ) which are internal indicators of scene contrast, retrieval convergence, and disagreement between various steps of the retrieval, all of which we have found to correlate with retrieval quality [3]. In Version 5, the case-by-case values of each of the parameters whose thresholds were used in the Version 4 acceptance tests,  $Y_k$ , are used in the generation of error estimates of the individual retrieved parameters. Values of four other retrieval convergence tests are included as well. Error estimates for  $T(p)$  and  $T_{skin}$  are computed according to

$$\delta X_i = \sum_{k=1}^N M_{ik} Y_k \quad (6)$$

where  $\delta X_i$  is the error estimate of retrieved geophysical parameter  $X_i$ ,  $Y_k$  is the value of the  $k^{\text{th}}$  test,  $M$  is a matrix with different values over ocean and land, and  $N$  is the number of tests used to determine the error estimate. Error estimates are, by definition, all positive. Three of the tests include AMSU-A observations. These tests are not used in Version 5 AO. Therefore,  $N=16$  in Version 5 and  $N=13$  in Version 5 AO. Other than the number of tests used, error estimates are computed and used in an analogous manner in Version 5 and Version 5 AO.

The error estimate for total precipitable water is computed in an analogous manner to that in equation 6, but is computed in terms of fractional error estimate

$$\delta \left( \frac{W_{tot} - W_{tot}^{truth}}{W_{tot}} \right) = \delta FE = \sum_{k=1}^N M_{ik} Y_k \quad (7)$$

where FE is the fractional error in total precipitable water. The error estimate for  $W_{tot}$  is obtained according to  $\delta W_{tot} = \delta FE(W_{tot})$ .

Total precipitable water is not derived directly in the physical retrieval. Rather, a water vapor profile retrieval is performed to give  $q(p)$  [2,3].  $W_{tot}$  is computed as the vertical integral  $\int_0^{p_{surf}} q(p) dp$ . Level by level error estimates for  $q(p)$  and for channel by channel clear column radiances  $\hat{R}_i$ , are generated in a slightly different manner, based on the values of the temperature profile error estimates themselves. More details about the generation of error estimates for all retrieved quantities and clear column radiances will be given in a separate publication.

### C. Determination of $M$

If one knows the actual errors, given by  $X_i - X_i^{truth}$ , the matrix  $M$  is determined in a straightforward manner, by finding  $M$  such that  $M$  minimizes the RMS difference of  $(\Delta X_i - \delta X_i)$ , where  $\Delta X_i = |X_i - X_i^{truth}|$ . In order to generate  $M$ , we used  $X_i$  and  $Y_k$  for all accepted Version 5 retrievals (that is all cases passing the Version 4 Stratospheric Temperature Test) on September 29, 2004, and used the colocated ECMWF 3 hour forecast as truth. The set of accepted cases used for training (Stratosphere good) is not the same for Version 5 and Version 5 AO, nor are the resulting coefficients. The coefficients  $M$  for each version are determined separately and used once and for all. The accuracy of sample error estimates, and their use for quality control, is shown later for global Version 5 and Version 5 AO retrievals run on January 25, 2003. Quality controlled retrieval results obtained for Version 4, using Version 4 quality control, are also shown for comparison.

### Surface skin temperature Quality Control

Ocean surface skin temperatures are measured very accurately from other EOS instruments such as MODIS and AMSR-E. AIRS ocean skin temperatures must be very accurate in order to provide additional useful information regarding sea surface temperature anomalies. In Version 4, sea surface temperatures were classified according to their ability to pass either a Tight SST test (best quality) or a Standard SST test (good quality). Monthly mean products were generated by including all cases passing the standard SST test. In Version 5, the ocean skin temperature error estimate  $\delta T_{skin}$  is used directly for quality control. Version 5 classifies ocean skin temperatures as good quality if  $\delta T_{skin} < 1.0K$  and best quality if  $\delta T_{skin} < 0.8K$ .

### D. Temperature profile Quality Control

As with surface skin temperature, case-by-case level-by-level error estimates for temperature profiles are also obtained using equation 6. These error estimates are subsequently used to determine a case-by-case characteristic pressure  $p_{best}$ , down to which the profile is considered of highest quality and acceptable for use for data assimilation purposes as well as for process studies. All accepted IR/MW profiles, that is, all cases passing the Version 4 Stratospheric Temperature Test, are assigned to have high quality down to at least 70 mb. The characteristic pressure  $p_{best}$  is defined as the largest pressure (somewhere between 70 mb and  $p_{surf}$ ) at which the error estimate in each of the next 3 pressure levels is not greater than a pressure dependent error estimate threshold  $\delta T(p)$ . These pressure dependent thresholds vary between 3.0K and 1.25K throughout the atmosphere, and are specified separately for land and ocean, each with different thresholds in Version 5 and Version 5 AO.

Pressure dependent thresholds are determined from a set of 3 threshold parameters  $\delta T_{70}$ ,  $\delta T_{mid}$ , and  $\delta T_{surf}$ , representative of error thresholds for  $T(p)$  at  $p = 70$  mb, at  $p =$

$P_{\text{surf}}/2$ , and at  $p = P_{\text{surf}}$  where  $P_{\text{surf}}$  is the surface pressure. The thresholds  $\delta T(p)$  at intermediate pressures are linearly interpolated in  $\log p$  between the given values. We have found it advantageous to have separate error thresholds for non-frozen ocean on the one hand, and land and ice on the other. Table I shows the Version 5 and Version 5 AO thresholds used at the Goddard DISC for both non-frozen ocean (called ocean) and other than non-frozen ocean (called land).

Table I  
Temperature Profile Thresholds (K)

	Ocean			Land		
	$\delta T_{70}$	$\delta T_{\text{mid}}$	$\delta T_{\text{surf}}$	$\delta T_{70}$	$\delta T_{\text{mid}}$	$\delta T_{\text{surf}}$
Version 5	1.75	1.25	2.25	2.25	2.0	2.0
Version 5AO	1.75	1.0	1.75	1.75	1.0	2.0

## VI. RESULTS FOR ONE FULL DAY

### A. Sea surface temperature error estimates and Quality Control

Fig. 2a shows the spatial distribution of all Version 5 non-frozen ocean surface skin temperature (SST) errors for ascending (daytime) orbits on January 25, 2003 for all cases in which the final IR/MW retrievals were generated. The predicted SST errors, obtained from Equation 6, are shown in Fig. 2b. Predicted errors are always positive, indicating the magnitude (but not the sign) of the uncertainty of a geophysical parameter. Fig. 2d shows the difference between the predicted error and the absolute value of the difference from the ECMWF 3 hour forecast. The largest errors are somewhat under-predicted and the smallest errors over-predicted, but the spatial correlation (0.76) between observed and estimated errors is very good.

Fig. 2c shows the spatial distribution of the errors for those sea surface temperatures for which the error estimate is less than 1K. These are classified as good quality and only these cases are used in the generation of the Version 5 monthly mean Level 3 SST product. The large negative SST bias shown in Fig. 2a is eliminated when the Standard Quality Control is used, and the spatial standard deviation of quality controlled SST errors is reduced from 4.46K to 0.84K.

Fig. 3 shows histograms of the counts of errors of quality controlled ocean SST retrievals as a function of SST differences from ECMWF “truth” on January 25, 2003, obtained using Version 4, Version 5, and Version 5 AO. The results shown for Version 5 and Version 5 AO use the standard quality test,  $\delta T_{\text{skin}} < 1.0\text{K}$ . The mean difference from ECMWF, the spatial standard deviation of the difference, the percent of all cases accepted, and the percent outliers (errors more than 3K from the mean) are indicated in Fig. 3. Version 5 has a higher yield and somewhat better accuracy than Version 5 AO, but the degradation obtained from

potential loss of the AMSU instrument is not great. Both Version 5 and Version 5 AO have higher yields than obtained using Version 4 with the Standard Criteria SST quality control, with a significantly greater accuracy as well. Version 4 Tight Criteria SST has very high accuracy but the yield is not high enough for climate use.

### B. Atmospheric temperature profile error estimates and Quality Control

Fig. 4a shows the differences of retrieved 300 mb temperatures from ECMWF “truth” for all accepted ascending orbit cases in January 25, 2003 using Version 5. Gray means missing data. This can be a result of orbit gaps, a missing granule (over central Africa), or (generally very cloudy) areas where successful retrievals were not performed (such as off the northwest coast of the U.S.). The area weighted global mean of the error without quality control is -0.17K, and its spatial standard deviation is 1.40K. Fig. 4b shows the predicted 300 mb errors, and Fig. 4d shows the differences between the predicted error and the absolute value of the actual error. The spatial correlation of predicted and absolute errors is 0.43, and the spatial standard deviation of the error in the prediction is 0.89K, showing reasonable skill between the actual “error” (which may itself be incorrect due to errors in the truth), and the predicted error.

Fig. 4c shows the 300 mb error of the quality controlled cases, i.e., cases in which  $p_{\text{best}} \geq 300\text{mb}$ . The spatial distribution of accepted cases is quite extensive, and the standard deviation of the errors for accepted cases has dropped to 1.18K. The largest differences from ECMWF for the accepted cases occur over Antarctica, Greenland, and Northern Siberia, in locations where the error estimates are low. These are regions in which the ECMWF “truth” may be of poorer quality and actual errors may be less than the errors shown in Fig. 4c.

Fig. 5 shows RMS differences of 1 km layer mean temperatures from ECMWF “truth” for global quality controlled Version 4, Version 5, and Version 5 AO retrievals, as well as the percent of all cases included in each set of statistics for each layer. The percent accepted at 70 mb represents the percent of all cases in which successful IR/MW retrievals were produced in each system. Version 5 and Version 5 AO both have a much higher yield beneath 200 mb than Version 4. Version 5 has comparable or better accuracy than Version 4 beneath 200 mb. Increasing spatial coverage of high quality retrievals is very important for both data assimilation purposes and climate and process studies. The improvement in Version 5 over Version 4 near the surface in both yield and accuracy is particularly noteworthy. This results primarily from improved performance over land. Both Version 5 and Version 5 AO have considerably lower errors than Version 4 between 100 mb and 200 mb, with a slightly lower yield. This is because not all cases passing the Version 4 Stratospheric Temperature Test were actually good all the way down to 200 mb. Version 5 performs somewhat better than Version 5 AO with regard to both yield and accuracy, but

Version 5 AO retrievals are still very good and provide an excellent backup mode should AMSU-A fail.

Fig. 6a shows the Version 5 retrieved cloud parameters for the ascending orbits of January 25, 2003. Different colors indicate cloud height (reds and purples are high clouds, blues and greens are mid-level clouds, and yellows and oranges are low clouds). Darker shades of the colors indicate larger cloud amounts. Figs. 6b-d show the spatial distribution of Quality Controlled retrieved 700 mb temperatures for Version 5 (Fig. 6b), Version 4 (Fig. 6c), and Version 5 AO (Fig. 6d). Data for areas in which the surface pressure is less than 700 mb are not included in Figs. 6b-6d. It is important to note that the spatial coverage of Quality Controlled 700 mb temperatures over land and at high latitudes is significantly higher in both Version 5 and Version 5 AO than in Version 4. Version 4 also rejects lower tropospheric temperatures in sunglint areas over ocean, while Version 5 does not.

Fig. 7 shows in gray the number of cases for each retrieved effective fractional cloud cover, in 0.5 percent bins, for the whole day January 25, 2003, as determined using Version 5. The effective fractional cloud cover is given by the product of the fraction of the field of view covered by clouds and the cloud emissivity at 11  $\mu\text{m}$ . The average global effective cloudiness was determined to be 44.5 percent in Version 5. There are peaks at 0 percent and 100 percent effective cloud cover, with a very smooth distribution at intermediate effective cloud fractions. The discontinuity at 90 percent cloud cover is an artifact arising from the switch from clouds retrieved primarily using the IR/MW retrieved state to clouds retrieved using the MW/strat IR state. Also shown, in different colors, is the percent of accepted retrievals as a function of retrieved effective cloud cover for Version 5 cases with  $p_{\text{best}} > 70$  mb,  $\geq 500$  mb, and  $\geq 700$  mb, as well as for ocean cases passing the Standard SST test. Almost all cases with retrieved effective cloud fraction less than 90 percent pass the Version 5 Stratospheric Temperature Test, which is equivalent to cases with  $p_{\text{best}} > 70$  mb. The percentage of accepted IR/MW retrievals with  $p_{\text{best}} > 70$  mb falls slowly with increasing cloud cover, from close to 100 percent at low cloud fractions to about 76 percent at close to 90 percent effective cloud cover. 79.6 percent of the global cases pass the Version 5 Stratospheric Temperature Test, with an average effective cloud fraction of 36.3 percent. 48.5 percent of the global cases have  $p_{\text{best}} \geq 500$  mb, with an acceptance rate of about 95 percent for low effective cloud fraction, falling to about 45 percent at 80 percent effective cloud fraction, and 28 percent at 90 percent effective cloud fraction. The mean effective cloud fraction for all Version 5 cases with  $p_{\text{best}} \geq 500$  mb is 31.2 percent. Only 26.3 percent of the cases have  $p_{\text{best}} \geq 700$  mb, primarily over ocean, with an acceptance rate near 85 percent for low cloud fractions and falling to 40 percent at 80 percent effective cloud fraction and 25 percent at 90 percent effective cloud fraction, and with an average effective cloud fraction of 32.8 percent.

The ensemble of cases used in the statistics for cases with  $p_{\text{best}} \geq 700$  mb is not the same as for the previous cases because all cases with  $p_{\text{surf}} \geq 700$  mb are excluded from these statistics. It is for this reason that the average cloud fraction for cases with  $p_{\text{best}} \geq 700$  mb can be higher than for cases with  $p_{\text{best}} \geq 500$  mb, even though the percentage of accepted cases as a function of increasing cloud cover is lower for  $p_{\text{best}} \geq 700$  mb than for  $p_{\text{best}} \geq 500$  mb.

The ensemble of cases used for statistics related to the cases passing the Standard SST criterion is again different as it includes non-frozen ocean cases only. The distribution of errors in these cases is given in Fig. 3. At low cloud fractions, roughly 75% of ocean cases pass the Version 5 Standard SST test, but only 15% of the ocean cases with 50% cloud cover pass the Standard SST test. The average cloud cover for all ocean cases passing the Standard SST test is 15.8%. The average cloud fractions for Version 5 cases passing the different criteria listed above are included in the first line of Table II.

Equivalent curves are shown with dashed lines for Version 4 cases passing the analogous tests: the Stratospheric Temperature Test, the Mid-Tropospheric Temperature Test, and the Lower Tropospheric Temperature Test, and Standard SST Test, respectively. The percent of cases accepted using a given criterion in Version 4 is always lower than in Version 5, and also falls more quickly as a function of increasing cloudiness. The improved ability of Version 5 to perform more accurate soundings under partial cloud cover is a direct result of using tropospheric sounding 15  $\mu\text{m}$   $\text{CO}_2$  channels only for the generation of clear column radiances  $\hat{R}_i$  for all channels, and using only 4.3  $\mu\text{m}$   $\text{CO}_2$  channels for tropospheric temperature sounding purposes. The average cloud fractions for Version 4 cases passing the analogous criteria to those of Version 5 are given in the second line of Table II. These average cloud fractions are considerably lower than those of Version 5. This confirms that the reason percent yields of Version 5 are larger than those of Version 4 (see Figs. 3 and 5) is that Version 5 produces more acceptable retrievals under cloudier cases. Fig. 7 includes analogous curves for Version 5 AO shown in solid lines with paler colors than used for Version 5. In general, Version 5 AO accepts somewhat fewer cases as a function of increasing cloud cover than Version 5, but still considerably more than Version 4. Average cloud fractions for Version 5 AO cases passing different criteria are included in the third line of Table II.

Table II  
Average Cloud Fraction of Accepted Cases

	All Cases	Stratosphere Good	500mb Good	700mb Good	Standard SST
Version 5	44.5	36.3	31.2	32.8	15.8
Version 4	44.3	34.8	26.8	21.0	10.2
Version 5AO	41.4	37.0	29.9	29.8	14.6

Fig. 8 shows the RMS difference between retrieved 1 km tropospheric layer mean temperatures and the collocated ECMWF 3 hour forecast for all accepted cases as a function of retrieved effective cloud fraction for both Version 5 and Version 4. Results are shown for four selected 1 km layers of the atmosphere. Only those cases passing the appropriate temperature profile tests are included in the statistics. Agreement degrades with increasing cloud cover, but only very slowly. The largest errors are in the 2 lowest layers in the atmosphere, at moderate to high cloud fraction, where the percentage acceptance rate is low. Errors in Version 5 as a function of cloud fraction are comparable to, or better than, those in Version 4 even though many more cloud cases are accepted in Version 5.

Fig. 8 also contains RMS errors of Standard Quality Controlled Sea Surface Temperature differences from those found in the ECMWF analysis as a function of cloudiness. These results are shown only up to 70% cloudiness, because the percentage of SST cases accepted at higher cloud fractions is extremely low, especially in Version 4. Quality controlled Version 5 SST results are more accurate than those of Version 4, and the accuracy degrades significantly more slowly with increasing cloud cover than in Version 4, even though the percent of SST cases accepted in Version 5 is considerably greater than in Version 4, especially at larger cloud fractions.

### C. Total precipitable water Quality Control: The Constituent Test

The error estimate for  $W_{\text{tot}}$  is obtained according to  $\delta W_{\text{tot}} = \delta \text{FE}(W_{\text{tot}})$ . This error estimate is also used in the Constituent Test, which provides Quality Control for the  $\text{H}_2\text{O}$ ,  $\text{O}_3$ ,  $\text{CO}$ , and  $\text{CH}_4$  profiles. These constituent profiles are flagged as unacceptable if  $\delta \text{FE} > 0.35$ . Other tests must be passed as well, in a manner described later.

Fig. 9a shows the error in total precipitable water vs. ECMWF,  $W_{\text{tot}} - W_{\text{tot}}^{\text{truth}}$ , for all Version 5 cases in which a successful IR/MW retrieval was performed. The global mean error is -0.07 cm (dry bias) and the standard deviation is 0.44 cm. Fig. 9b shows the predicted fractional error  $\delta \text{FE}$  for these cases, and Fig. 9d shows the difference between the absolute value of the actual error and the predicted error, given by the product of  $\delta \text{FE}$  and  $W_{\text{tot}}$ , for all cases. The correlation of observed and predicted errors is 0.52, with very little bias. Fig. 9c shows the error in quality controlled total precipitable water, that is, cases in which  $\delta \text{FE} \leq 0.35$ . The spatial coverage of quality controlled precipitable water is very good, and the standard deviation of the error has dropped to 0.39 cm.

Fig 10 shows RMS percent differences of quality controlled retrieved 1 km layer integrated precipitable water from those of ECMWF “truth”, as well as the percent yield of quality controlled retrievals for each layer in Version 5, Version 5 AO, and Version 4. In both Version 5 and Version 4, layer precipitable water for a given layer is flagged as good if both the Constituent Test is passed and the temperature profile at the bottom of the layer is also flagged as good. As

with temperature profiles, the accuracy of quality controlled Version 5 water vapor retrievals is comparable to or better than those of Version 4 at all levels, but with considerably greater spatial coverage. Version 5 AO water vapor retrievals are also of higher quality than those of Version 4, with greater spatial coverage as well. The fact that Version 5 AO retrievals are slightly degraded at the lowest levels compared to those of Version 5 and Version 4 is a result of the loss of the benefit of use of AMSU A channels 1, 2, and 15 in the water vapor retrieval step. These AMSU A channels are sensitive to boundary layer water vapor, especially over ocean. The HSB 183 GHz water vapor sounding channels are not used in any of the retrievals because the HSB instrument failed early in the mission.

### D. The ability to generate accurate Quality Controlled AIRS Only retrievals

The overall approach used in the AIRS Only retrieval system is completely analogous to that used in the AIRS/AMSU retrieval system. The procedure starts with the generation of an initial state  $X^0$  consistent with observed radiances. This state is needed so as to derive the initial set of cloud cleared radiances  $\hat{R}_1^0$ , which in turn are used to generate the regression state  $X^{\text{reg}}$ . Once  $X^{\text{reg}}$  is obtained, physical retrievals can be performed in a completely analogous manner, using only AIRS cloud cleared radiances, or AIRS cloud cleared radiances in conjunction with AMSU radiances.

In Version 4 and earlier Versions,  $X^0$  was taken as a state consistent with observed AMSU radiances so as to be unbiased in the vertical. AMSU radiances were considered to be an indispensable part of the sounding system as they were critical to the ability to produce values of  $R_1^0$  that are accurate enough for use in the generation  $X^{\text{reg}}$ . This approach was not feasible for use in an AIRS Only system. Therefore, the alternative approach of obtaining  $X^0$  using a cloudy regression with observed AIRS radiances was developed. Without this development, generation of AIRS only retrievals would not have been possible within our basic retrieval methodology which uses  $\hat{R}_1$  to produce accurate quality controlled soundings under particularly cloudy conditions. This “fallback” approach used to generate the startup state  $X^0$  performed extremely well under most cloud conditions. Under some very cloudy conditions, the state  $X^0$  can be very poor. Such cases are easily detected by the improved Quality Control methodology used in Version 5.

Figs. 3, 5, and 10 show that Quality Controlled Version 5 AO retrievals are almost comparable to those of Version 5, in terms of both accuracy and yield, and are superior to those of Version 4, which used both AIRS and AMSU observations. The ability to perform accurate Quality Controlled AIRS Only retrievals would not have been possible without the two major improvements in Version 5 physical retrieval methodology compared to Version 4: the ability to retrieve  $T(p)$  using primarily shortwave  $\text{CO}_2$  sounding channels while using longwave tropospheric sounding  $\text{CO}_2$  channels to generate



clear column radiances  $\hat{R}_i$ ; and the ability to generate accurate case-by-case level-by-level error estimates for use in Quality Control.

Fig. 11 shows global RMS errors, compared to ECMWF “truth”, of 1 km layer mean temperatures obtained in the different steps of the Version 5 AO retrieval process, with Quality Control (solid lines) and without Quality Control (dashed lines), for all the cases passing the Stratospheric Temperature Test. Results for nighttime cases and daytime cases are shown separately.

The first thing to note is that Quality Controlled results during the day are comparable to, if not better than, those at night. This shows that the effect of non-LTE on the radiances is well accounted for in the physical retrieval system, as well as effects of solar radiation reflected by clouds and by the surface. Solar radiation reflected by clouds is accounted for indirectly in the generation of  $\hat{R}_i$ , and solar radiation reflected by the surface is accounted for directly because the surface spectral bi-directional reflectance is solved for as part of the retrieval process. The Quality Controlled physical retrievals improve on the Quality Controlled regression and cloudy regression both at night and during the day. The improvement is in fact greater during the day, because the physical retrieval can account for the effects of non-LTE and reflected solar radiation, while the two regressions cannot.

One may think, by comparing the solid lines, that the degree of improvement of the Quality Controlled physical retrieval over each regression is relatively small. This is misleading, however, because in Fig. 11, the same Quality Control is used for all steps in the retrieval process. The ability to perform this Quality Control is intrinsic to the physical retrieval process. A more reasonable comparison of the benefit of the physical retrieval over regression is to compare the solid black line with the dashed red or dashed blue lines because the relevant Quality Control could not be applied if only regressions were performed. Without the ability to do error estimates and use them for Quality Control, the AIRS Only approach, based on use of a cloudy regression, would not be practical.

## VII. GENERATION OF LEVEL 3 PRODUCTS

The Goddard DISC generates Level 3 products, averaged over space and time, on a  $1^\circ \times 1^\circ$  lat-long grid, averaged separately for ascending (1:30 p.m.) and descending (1:30 a.m.) orbits. These products are generated for daily, 8-day, and monthly mean time periods. The Quality Control used in the generation of the Level 3 products is related to, but not identical, to the Quality Control described in Section V. Atmospheric temperatures passing the Quality Control described in Section V are flagged as QC=0 in the level 2 data. QC=0 refers to products of the highest quality, recommended for data assimilation purposes.

Level 3 products are used primarily for climate purposes, as opposed to data assimilation purposes. Absolute accuracy

of individual soundings is less important for climate purposes than it is for data assimilation purposes. On the other hand, it is very important to have good spatial coverage in the Level 2 products used in the generation of monthly mean Level 3 products, and the Level 2 results should be unbiased. The temperature profile Quality Control methodology used in the statistics shown in Fig. 5 was designed for data assimilation purposes. Global spatial coverage of QC=0 retrievals at 300 mb is reasonably good for Version 4, as well as for Version 5 and Version 5 AO. The percentage of accepted retrievals, and the corresponding spatial coverage of QC=0 retrievals, drops off sharply lower in the atmosphere in all three versions, especially Version 4. Therefore, for climate purposes, some soundings other than those with QC=0 must be included in the generation of Level 3 products.

### A. Level 3 approach in Version 4

If only Version 4 QC=0 soundings at 700 mb, with spatial coverage shown in Fig. 6c, were used to generate level 3 Version 4 700 mb temperatures, a very poor monthly mean temperature product would be the result. Therefore, the level 3 Version 4 temperature profile products which were generated by the Goddard DISC use a relaxed Quality Control methodology. Temperatures 200 mb and above (at pressures less than 200 mb) were included in the generation of the DISC Version 4 Level 3 product if the Stratospheric Temperature Test was passed, while temperatures at all pressures beneath 200 mb were included in the generation of the Version 4 level 3 product if the Mid-Tropospheric Test was passed. The Version 4 Mid-Tropospheric Temperature Test is much less stringent than the Version 4 Lower-Tropospheric Test used for the 700 mb Quality Control, which results in the spatial coverage shown in Fig. 6c. Temperatures in the lowest 3 km of the atmosphere passing the Mid-Tropospheric Temperature Test, but not the Lower-Tropospheric Temperature Test, were flagged as QC=1 in Version 4, which means that they are suitable for use in climate studies, but not for use in data assimilation or process studies.

Fig. 12a is identical to Fig. 6c, and Fig. 12c is analogous to Fig. 6c, but shows the Version 4 700 mb spatial coverage for all cases passing the Version 4 Mid-Tropospheric Temperature Test, rather than the Lower-Tropospheric Temperature Test, that is, all Version 4 700 mb temperatures with QC=0 or 1. All Version 4 Level 3 products that were produced at the Goddard DISC were generated by averaging all retrieved quantities with QC=0 or QC=1. The spatial coverage of such products for Version 4 is shown in Fig. 12c. The word  $p_{\text{good}}$  at the DISC represents the pressure down to which temperatures are included in the generation of the Level 3 product.

### B. Level 3 approach used in Version 5

An analogous procedure is used in the generation of the DISC Version 5 Level 3 atmospheric temperature products. QC is set equal to 0 above  $p_{\text{best}}$  for all successful IR/MW retrievals. Over non-frozen ocean, spatial coverage of QC=0

retrievals at the surface is quite extensive and adequate for the generation of climate records. Therefore, over non-frozen ocean,  $p_{\text{good}}$  is set equal to  $p_{\text{best}}$  and no additional cases have  $QC=1$  to supplement the  $QC=0$  cases used in the generation of the Level 3 temperature profile products. Over land, ice, and coasts, spatial coverage for  $QC=0$  retrievals drops off considerably near the surface. In these locations, for  $p$  greater than 300 mb,  $p_{\text{good}}$  is set equal to  $p_{\text{surf}}$  and  $QC$  is set equal to 1 between  $p_{\text{best}}$  and  $p_{\text{surf}}$  as long as  $p_{\text{best}}$  is greater than 300 mb.  $QC$  is set equal to 2 at pressures in which  $QC$  is not 0 or 1, and at all pressures for all cases in which a successful IR/MW retrieval was not performed. Fig. 12b is the same as Fig. 6b, while Fig. 12d shows the spatial coverage of Version 5 retrievals for the  $QC=0$  or 1 cases which are included in the generation of Level 3 products at 700 mb. The spatial distributions shown in Figs. 12c and d are the same as used to generate level 3 products for all temperatures beneath the appropriate pressure for Version 4 (200 mb) or Version 5 (300 mb). Level 3 products were not generated for Version 5 AO retrievals.

The 1 day spatial coverage of Version 4 700 mb retrievals with  $QC=0$  or  $QC=1$  is considerably improved over that with only  $QC=0$ , but is still poorer than that of Version 5 or Version 5 AO. Even if everything else were comparable between Version 4 and Version 5 products, Version 5 would produce a better monthly mean Level 3 product as a result of this improved daily spatial sampling.

### C. Effect of different spatial sampling on the generation of interannual monthly mean differences

Fig. 13 demonstrates the effect of improved sampling of Version 5 compared to Version 4 in terms of the interannual monthly 700 mb temperature difference for January 2004 minus January 2003. The Version 5 interannual difference is shown in Fig. 13a and the Version 4 interannual difference is shown in Fig. 13b. Red indicates January 2004 was warmer than January 2003, and blue means January 2004 was cooler. Areas in which the surface pressure is less than 700 mb, such as over part of Antarctica and the Himalayas, are shown in gray. The area-averaged global mean interannual differences (excluding gray areas) are indicated in the figures.

The basic spatial patterns shown in Figs. 13a and 13b are very similar to each other. However, the features are stronger and more spatially coherent in Version 5 than in Version 4. Both figures show January 2004 to be colder globally at 700 mb than January 2003, but the amount of cooling found in Version 5 (-0.14K) was substantially less than that found in Version 4 (-0.28K).

One reason for these differences in the global mean interannual difference is a result of the different spatial sampling of accepted retrievals on a given day in Version 5 compared to Version 4. Another reason is differences between the Version 5 retrievals and Version 4 retrievals for a given sounding. Figs. 13c and 13d address the sampling issue

directly. Both of these figures show the interannual difference of the monthly mean values of the ECMWF 3 hour forecasts, with areas sampled identically to those cases accepted in Version 5 (Fig. 13a) and in Version 4 (Fig. 13b) respectively. Any differences between Figs. 13c and 13d are due only to sampling, because the ECMWF 700 mb temperatures being averaged are otherwise identical in the two figures. To first order, Figs. 13a and 13c, both sampled alike, look very much like each other, as do Figs. 13b and 13d. Therefore, the improved coherence and intensity of the patterns of interannual differences shown in Version 5, as compared to Version 4, are the result of the improved spatial sampling on a daily basis obtained in Version 5. The resultant global mean values of the interannual differences of ECMWF 700 mb temperature are also affected by daily sampling differences. The global mean interannual 700 mb temperature difference between January 2004 and January 2003 using ECMWF data sampled according to Version 5 retrievals, is given by -0.16K, and is probably more indicative of the true global mean interannual difference than is the value obtained using Version 4 sampling, -0.23K, because of more complete sampling in Version 5. The global mean interannual 700 mb temperature difference value shown in Fig. 13a agrees better to the estimate of truth shown in Fig. 13c, both sampled the same way, than do the global mean values shown in Figs. 13b and 13d, both also sampled the same way. This improved agreement with “truth” is a result of changes in the retrieval methodology and Quality Control used in Version 5 compared to Version 4.

Fig. 14 shows the global mean interannual differences of mandatory level temperatures between 850 mb and 70 mb obtained using Version 5 and Version 4, as well as obtained using appropriately sampled ECMWF data. Sampling differences between Version 4 and Version 5 become significant beneath 400 mb, as evidenced by the difference between the pink and red lines. In general, interannual global mean temperature differences determined from Version 5 observations (black) match those found in ECMWF (red) to better than 0.1K with the exception of 100 mb and 70 mb. Agreement of Version 4 interannual differences (gray) with ECMWF is poorer, especially when ECMWF is better sampled.

## VIII. AVAILABILITY OF AIRS SCIENCE TEAM PRODUCTS

AIRS/AMSU Version 5 Level 2 and Level 3 products are available at the Goddard DISC. The data can be found at <http://disc.sci.gsfc.nasa.gov/AIRS/data-holdings/by-data-product/>, and the documentation can be found at <http://disc.sci.gsfc.nasa.gov/AIRS/documentation>. Level 2 (spot by spot) data contain both the values of the retrieved products and their error estimates, as well as Quality Flags. Each retrieval also contains values of all the predictors used to generate the error estimates as well as the words  $p_{\text{best}}$  and  $p_{\text{good}}$ . The word  $p_{\text{best}}$  is as defined in this paper, and represents the pressure down to which it is recommended that the retrievals should be used for data assimilation and process studies. If

users generate their own accuracy and yield statistics for temperature and moisture profiles, soundings should only be used down to  $p_{\text{best}}$ , as done in the generation of Figs. 5 and 10. The DISC word  $p_{\text{good}}$  indicates the pressure down to which cases will be included in the generation of the Level 3 temperatures. DISC Level 3 products are composites of retrievals with QC=0 or QC=1, averaged over 1 deg by 1 deg latitude-longitude grid boxes and over 1 day, 8 day, and monthly mean time periods. Data from ascending (1:30 PM) and descending (1:30 AM) orbits are averaged separately. Fig. 12d is an example of the Version 5 700 mb Level 3 product shown for ascending orbits on January 25, 2003. Version 5 Level 3 constituent profile products for water and trace gases were generated using only those cases passing the Constituent Good Test.

#### REFERENCES

- [1] T. S., Pagano, H. H. Aumann, D. E. Hagan, and K. Overoye, Prelaunch and in-flight radiometric calibration of the Atmospheric Infrared Sounder (AIRS), *IEEE Trans. Geosci. Remote Sensing*, 41, 265-273, 2003.
- [2] J. Susskind, C. D. Barnett, and J. M. Blaisdell, Retrieval of atmospheric and surface parameters from AIRS/AMSU/HSB data in the presence of clouds, *IEEE Trans. Geosci. Remote Sensing*, 41, 390-409, 2003.
- [3] J. Susskind, C. Barnett, J. Blaisdell, L. Iredell, F. Keita, L. Kouvaris, G. Molnar, and M. Chahine, Accuracy of geophysical parameters derived from Atmospheric Infrared Sounder/Advanced Microwave Sounding Unit as a function of fractional cloud cover, *J. Geophys. Res.*, 111, D09S17, doi:10.1029/2005JD006272, 2006.
- [4] M. D. Goldberg, Y. Qu, L. M. McMillin, W. Wolf, L. Zhou, and M. Divakarla, AIRS near-real-time products and algorithms in support of operational numerical weather prediction, *IEEE Trans. Geosci. Remote Sensing*, 41, 379-389, 2003.
- [5] M. T. Chahine, Remote sensing of cloudy atmospheres. II. Multiple cloud formations, *J. Atmos. Sci.*, 34, 744-757, 1977.
- [6] P. W. Rosenkranz, Retrieval of temperature and moisture profiles from AMSU-A and AMSU-B measurements, in *Proc. IGARSS*, 2000.
- [7] M. T. Chahine, Remote sensing of cloudy atmospheres. I. The single cloud layer, *J. Atmos. Sci.*, 31, 233-243, 1974.
- [8] S. G. DeSouza-Machado, L. L. Strow, S. E. Hannon, H. E. Motteler, M. Lopez-Puertas, B. Funke, and D. P. Edwards, Fast forward radiative transfer modeling of 4.3  $\mu\text{m}$  non local thermodynamic equilibrium effects for infrared temperature sounders, *Geophys. Res. Letters*, 34, L01802, doi:10.1029/2006GL026684, 2007.
- [9] P. van Delst and X. Wu, A high resolution infrared sea surface emissivity database for satellite applications, Eleventh International TOVS Study Conf., Budapest, Hungary, 2000.
- [10] X. Wu and W. L. Smith, Emissivity of rough sea surface for 8-13  $\mu\text{m}$ : Modeling and verification, *Appl. Opt.*, 36, 2609-2619, 1997.

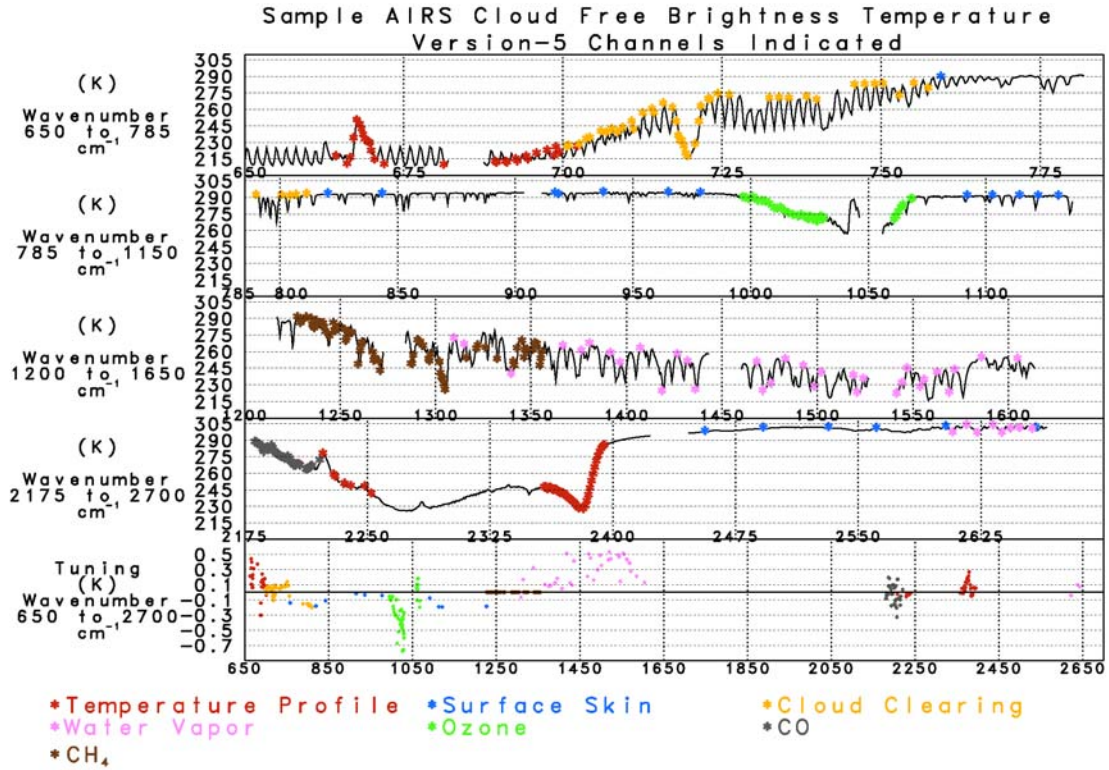


Fig. 1

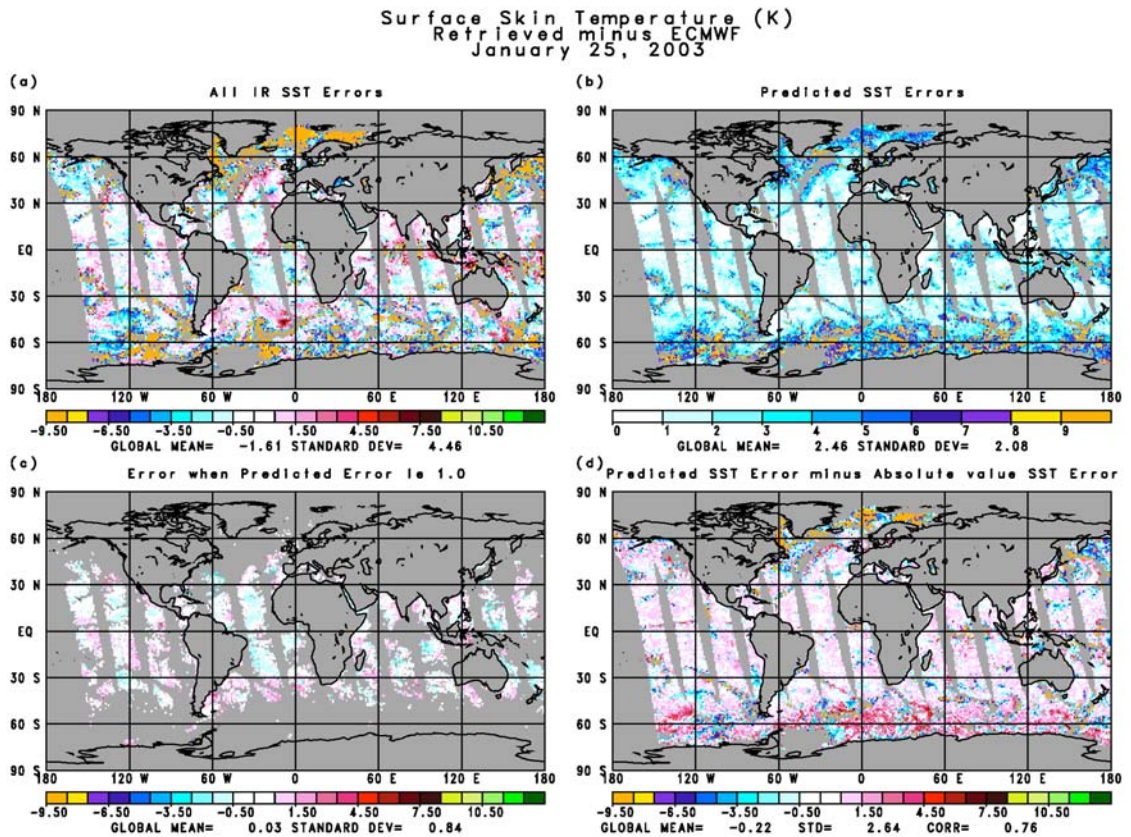


Fig. 2

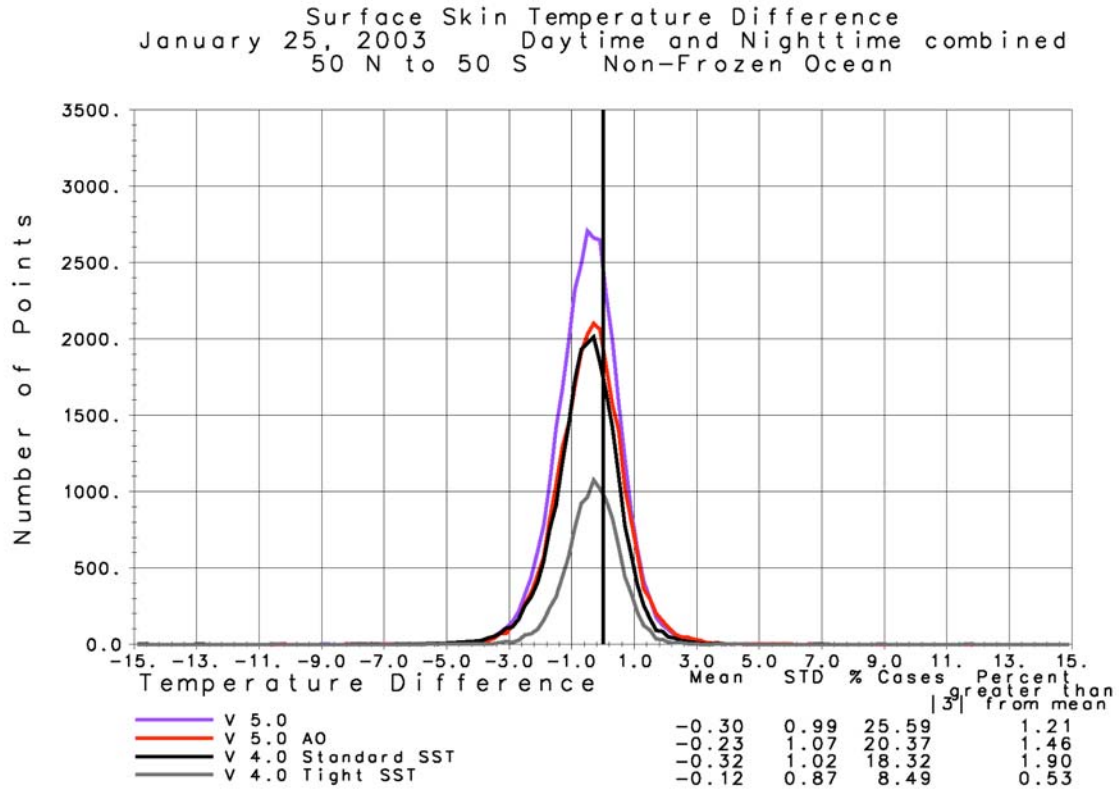


Fig. 3

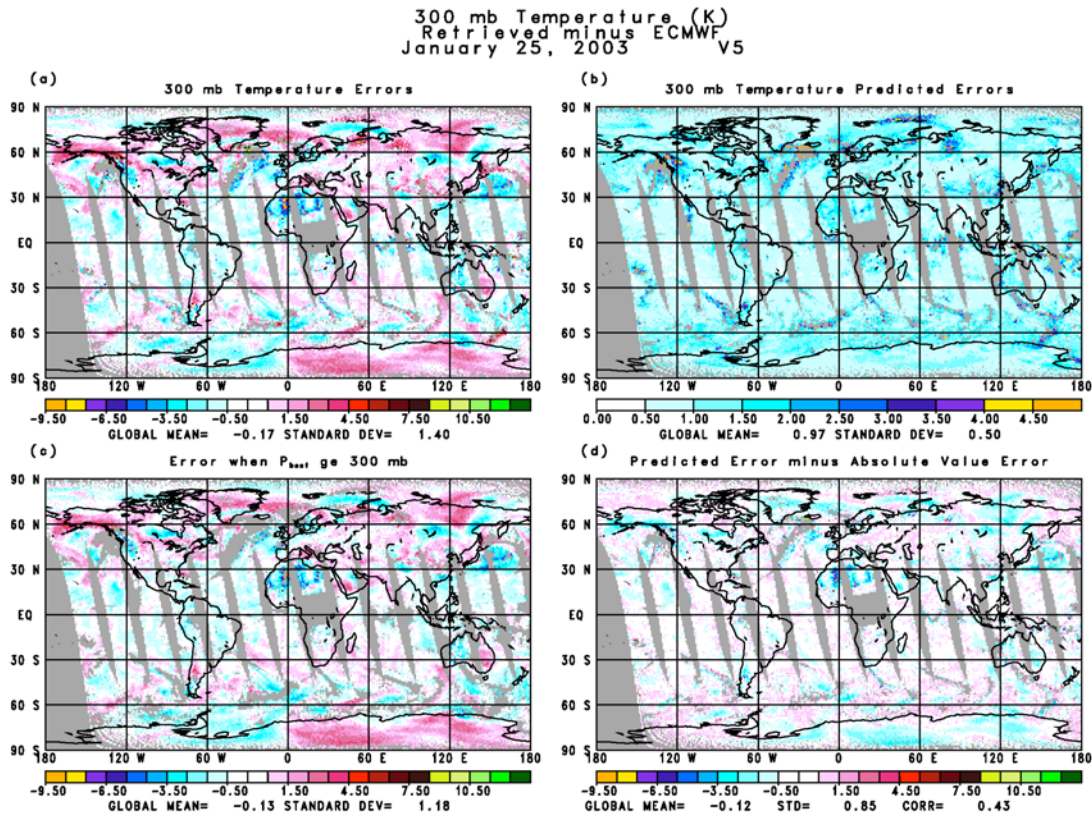


Fig. 4

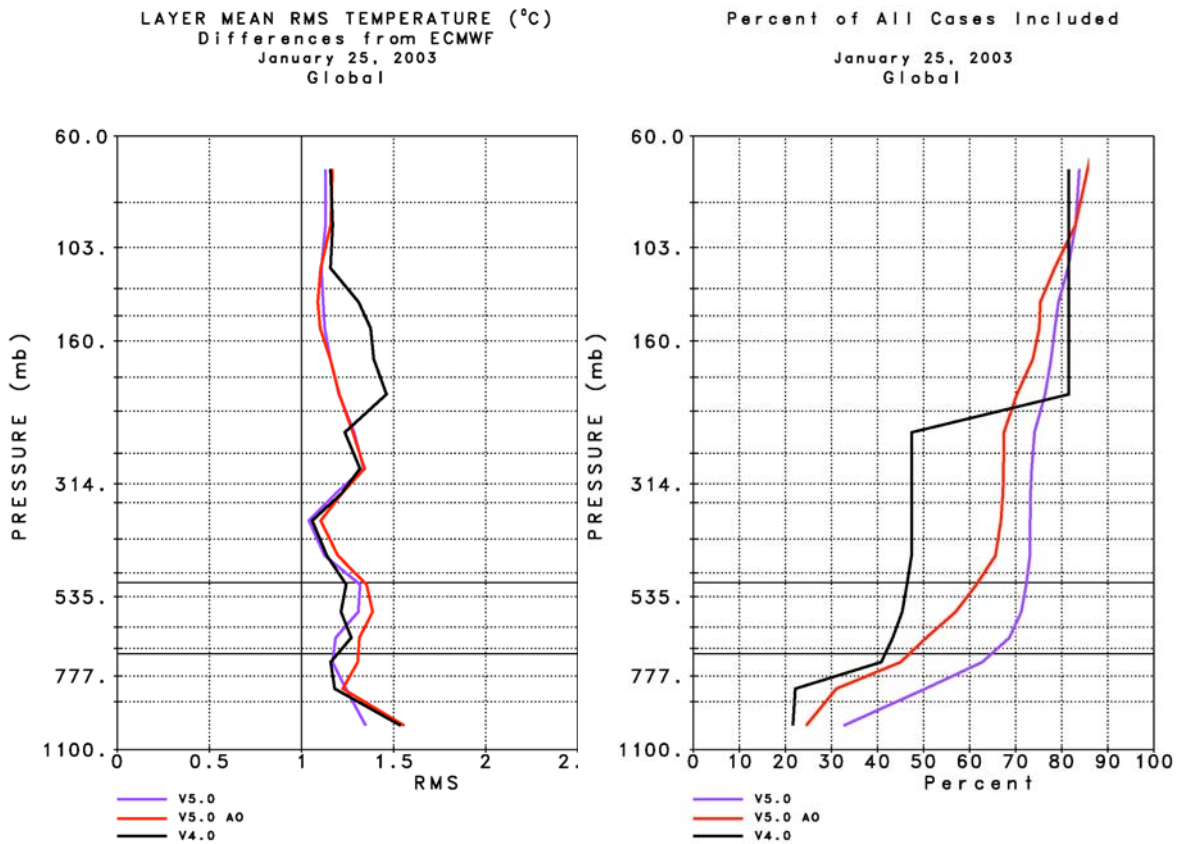


Fig. 5

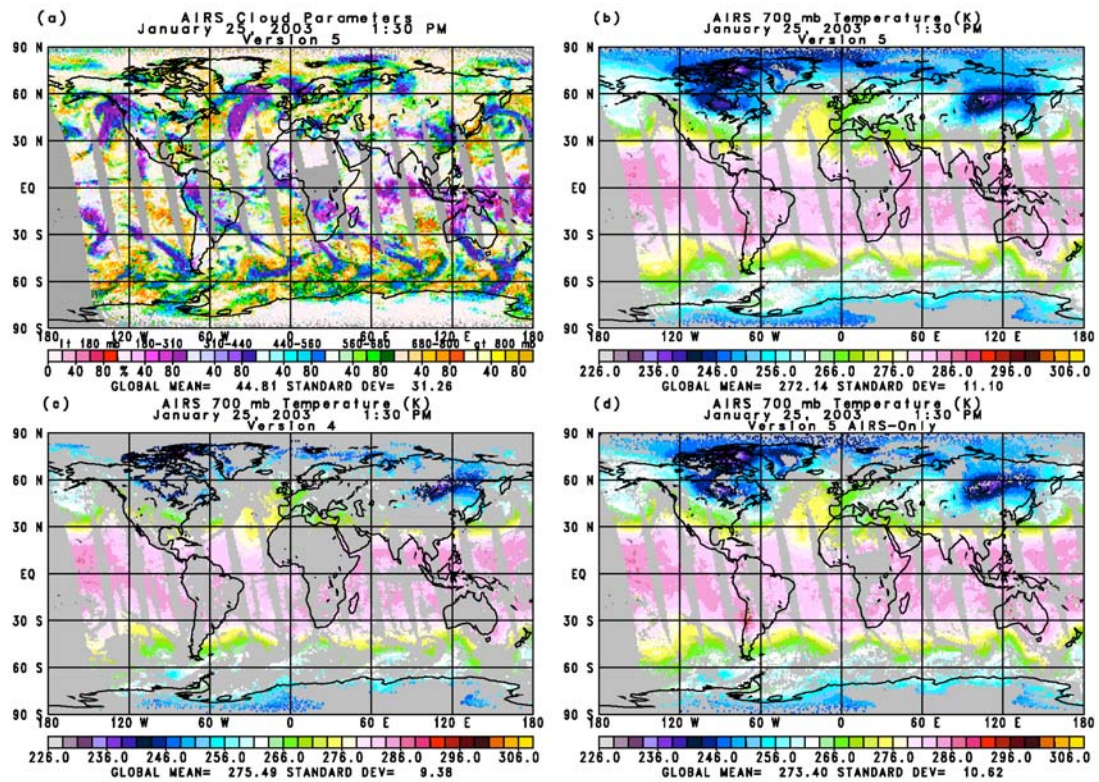


Fig. 6

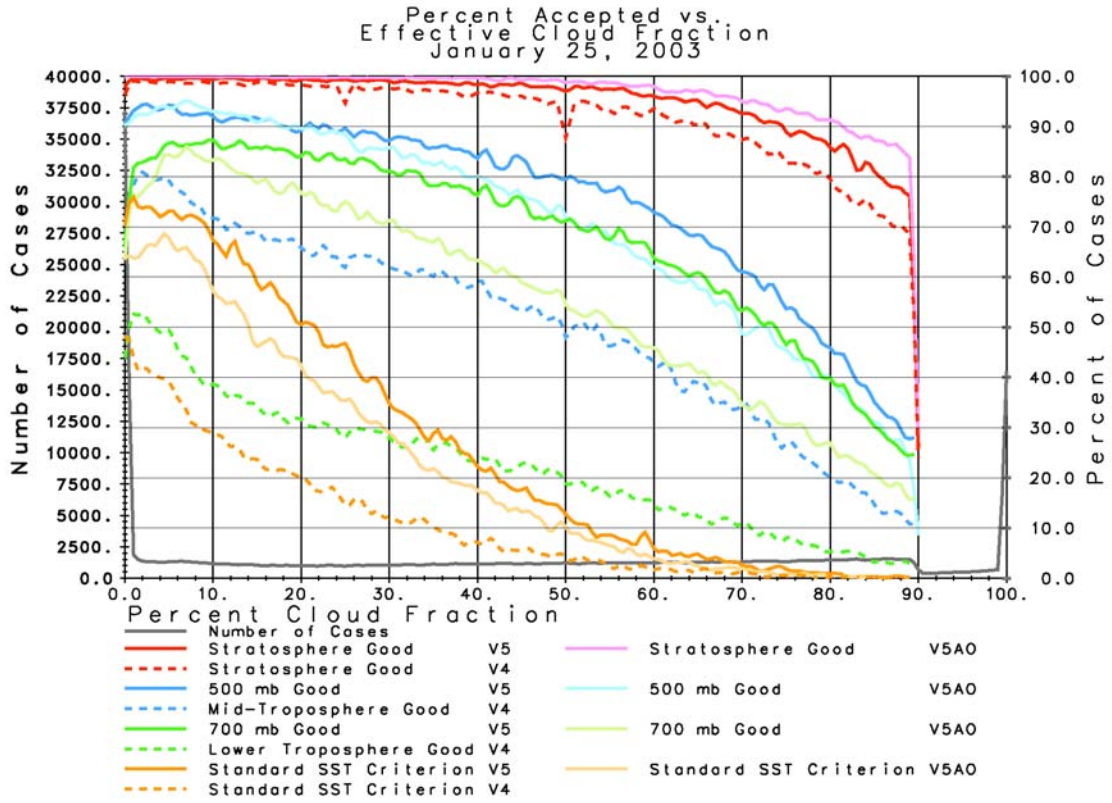


Fig. 7

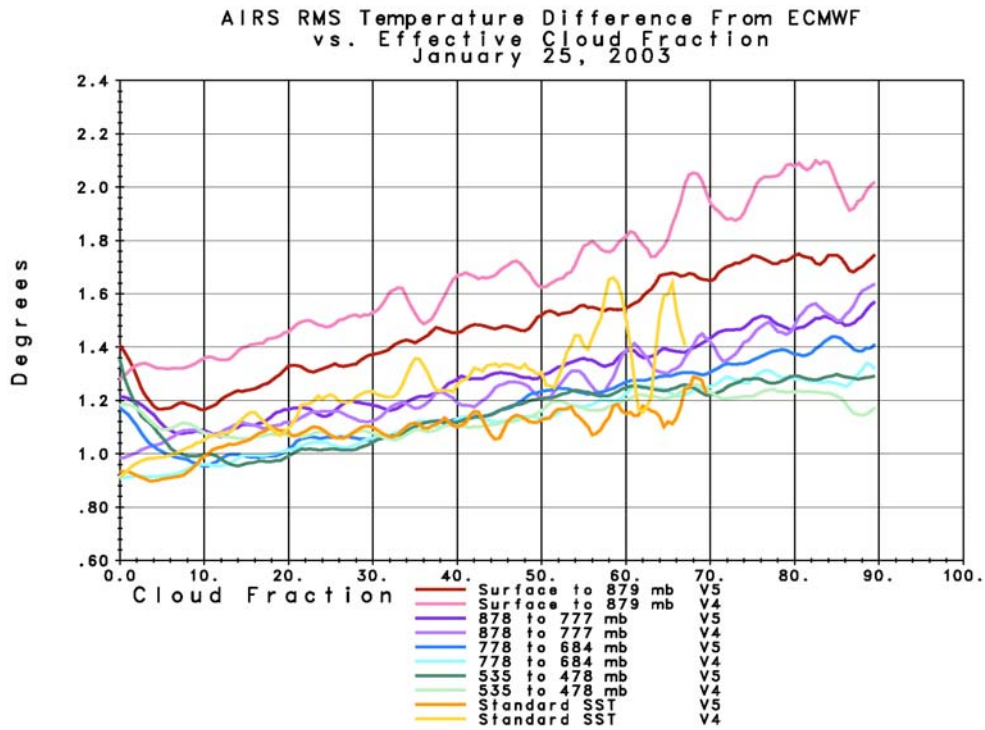


Fig. 8

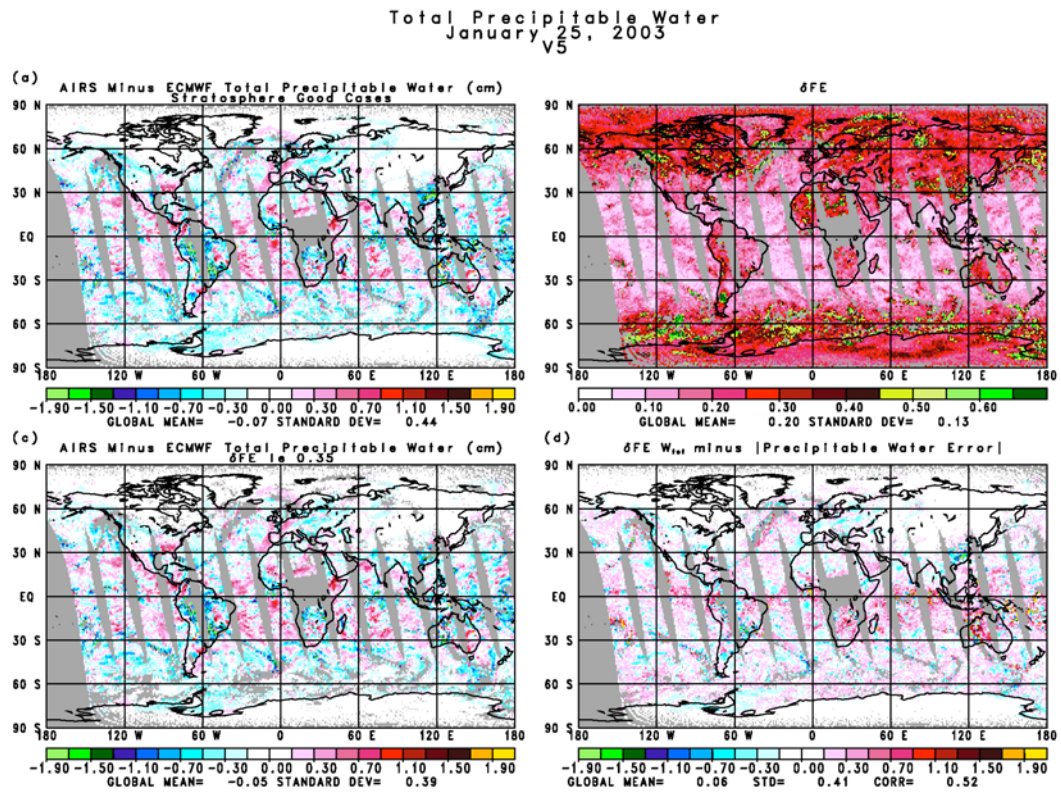


Fig. 9

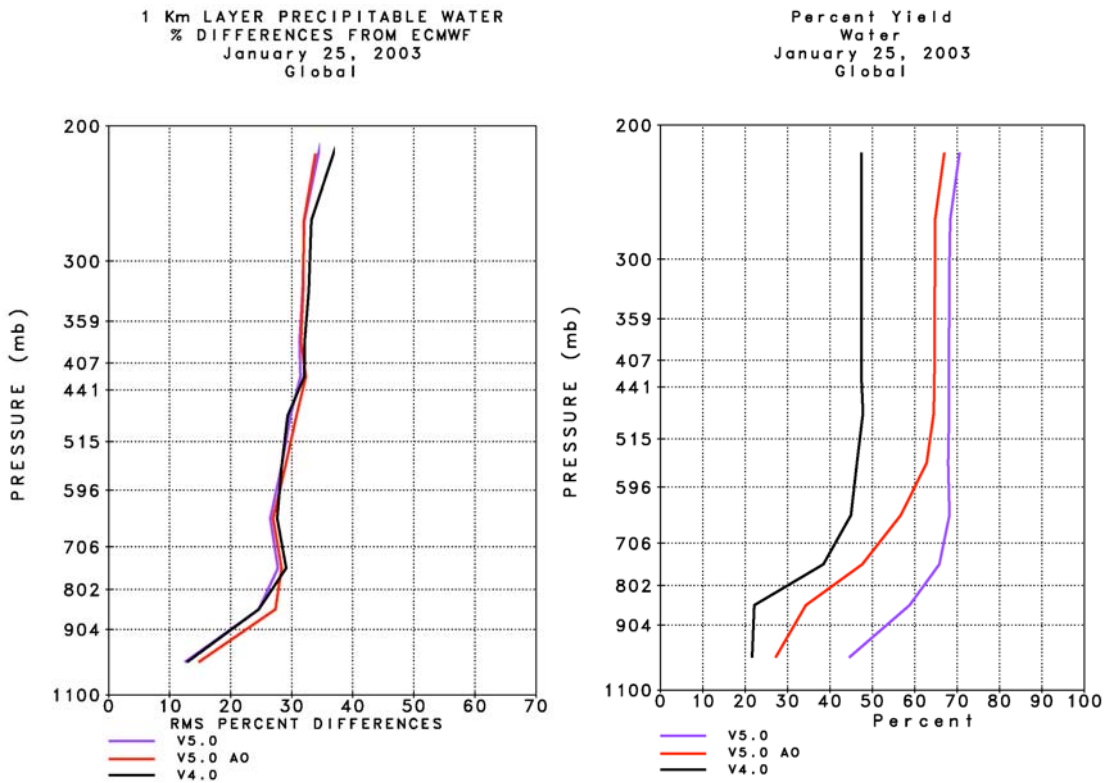


Fig. 10



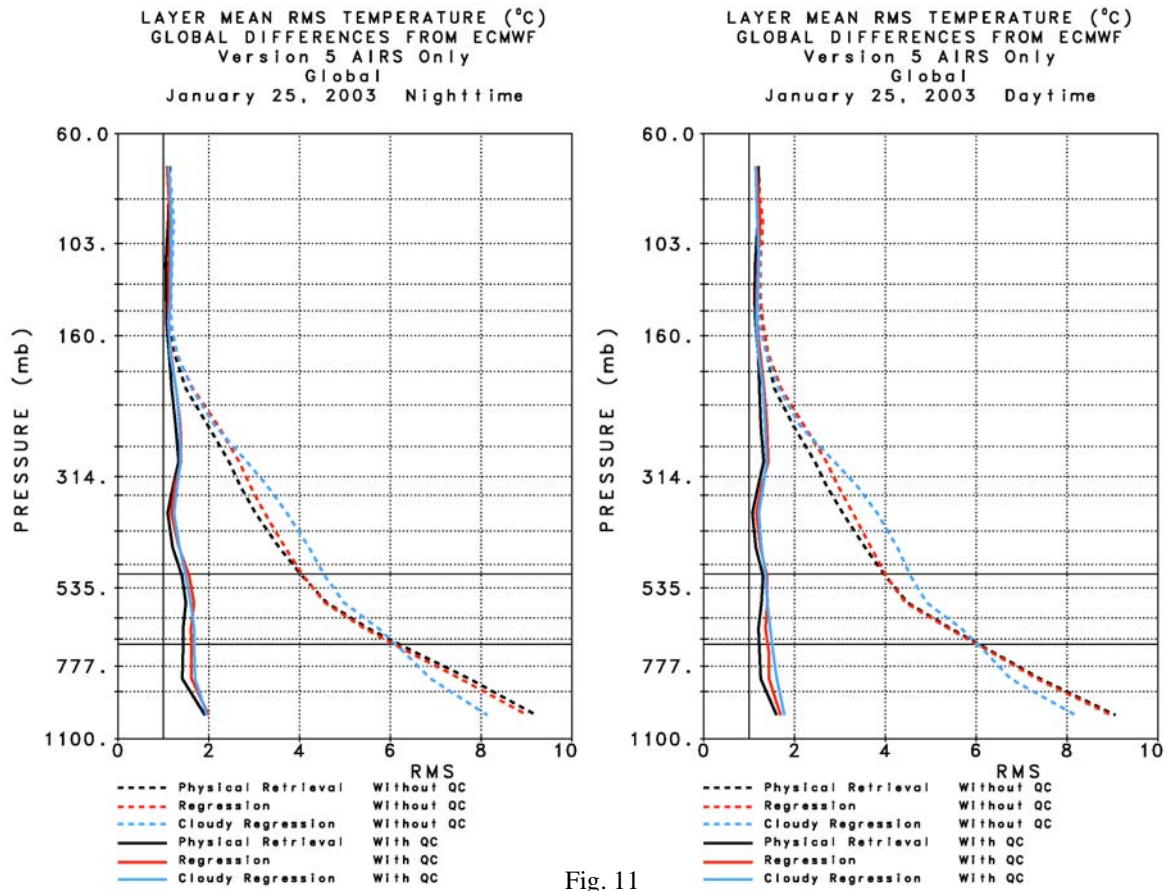


Fig. 11

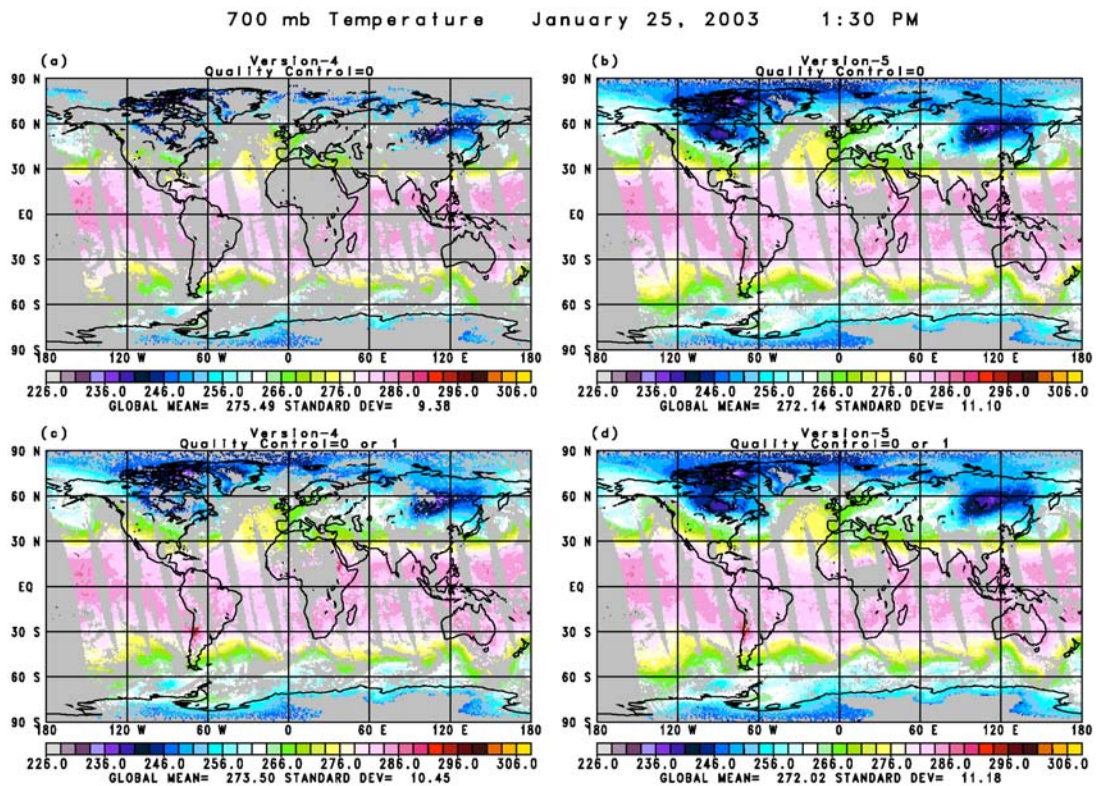


Fig. 12

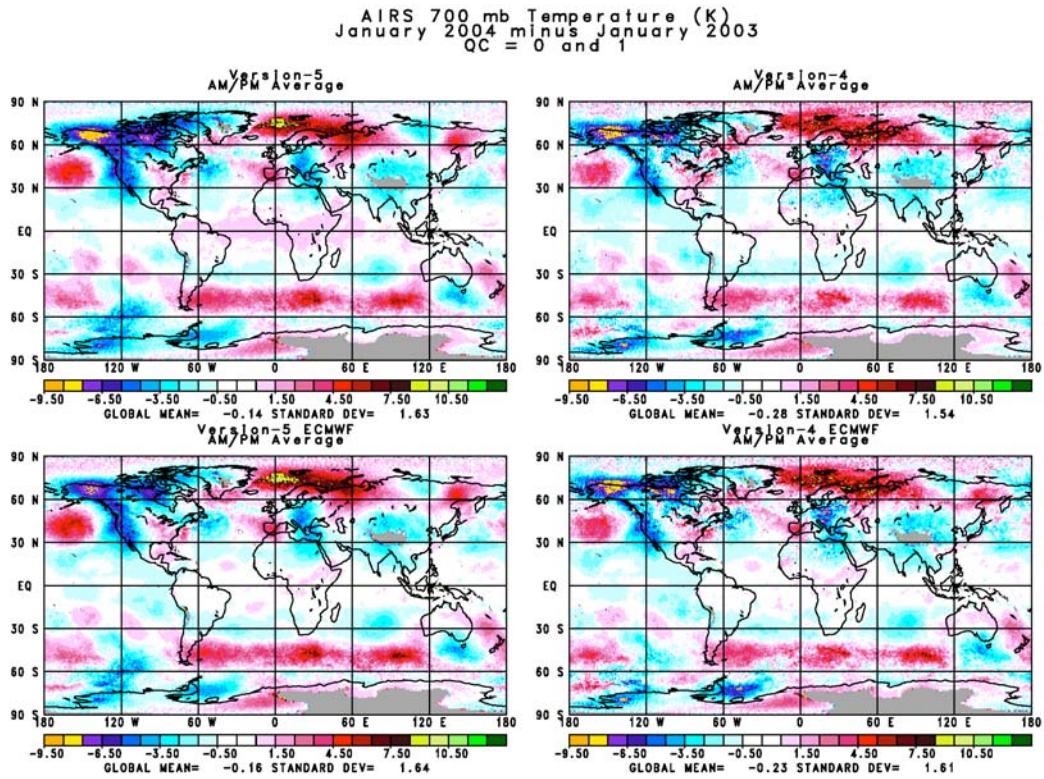


Fig. 13

Interannual Temperature Difference (°C)  
Global Mean  
January 2004 minus January 2003

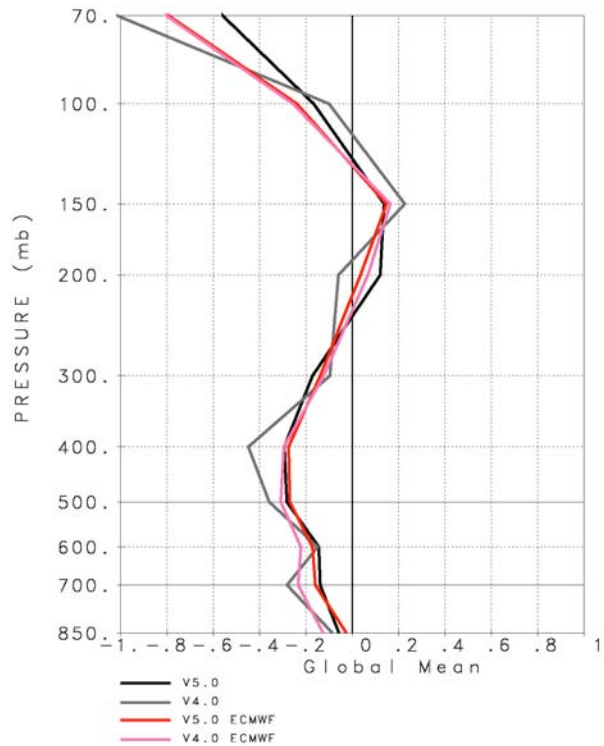


Fig. 14