

Accuracy of Geophysical Parameters Derived from AIRS/AMSU as a Function of Fractional Cloud Cover

Joel Susskind^a, Chris Barnet^b, John Blaisdell^c, Lena Iredell^c, Fricky Keita^c, Lou Kouvaris^c,
Gyula Molnar^d, and Moustafa Chahine^e

^aNASA Goddard Space Flight Center, Greenbelt, MD, USA 20771

^bNOAA/NESDIS, Camp Springs, MD, USA 20746

^cScience Applications International Corporation, NASA GSFC, Greenbelt, MD, USA 20771

^dJoint Center for Earth Systems Technology, NASA GSFC, Greenbelt, MD, USA 20771

^eJet Propulsion Laboratory, Pasadena, CA, USA 94709

Abstract. 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 system. The primary products of AIRS/AMSU are twice daily global fields of atmospheric temperature-humidity profiles, ozone profiles, sea/land surface skin temperature, and cloud related parameters including OLR. The sounding goals of AIRS are to produce 1 km tropospheric layer mean temperatures with an rms error of 1K, and layer precipitable water with an rms error of 20 percent, in cases with up to 80 percent effective cloud cover. The basic theory used to analyze AIRS/AMSU/HSB data in the presence of clouds, called the at-launch algorithm, was described previously. Pre-launch simulation studies using this algorithm indicated that these results should be achievable. Some modifications have been made to the at-launch retrieval algorithm as described in this paper. Sample fields of parameters retrieved from AIRS/AMSU/HSB data are presented and validated as a function of retrieved fractional cloud cover. As in simulation, the degradation of retrieval accuracy with increasing cloud cover is small and the RMS accuracy of lower tropospheric temperature retrieved with 80 percent cloud cover is about 0.5 K poorer than for clear cases. HSB failed in February 2003, and consequently HSB channel radiances are not used in the results shown in this paper. The AIRS/AMSU retrieval algorithm described in this paper, called Version 4, become operational at the Goddard DAAC (Distributed Active Archive Center) in April 2003 and is being used to analyze near-real time AIRS/AMSU data. Historical AIRS/AMSU data, going backwards from March 2005 through September 2002, is also being analyzed by the DAAC using the Version 4 algorithm.

1. Introduction

AIRS/AMSU/HSB (Atmospheric InfraRed Sounder/Advanced Microwave Sounding Unit/Humidity Sounder Brazil) is a state of the art advanced infra-red microwave sounding system that was launched on the EOS Aqua platform in a 1:30 AM/PM sun synchronous orbit on May 4, 2002. An overview of the AIRS instrument and the objectives for AIRS/AMSU/HSB is given in Aumann et al. (2003). The sounding goals of AIRS are to produce 1 km tropospheric layer mean temperatures with an rms error of 1K, and layer precipitable water with an rms error of 20 percent, in cases with up to 80 percent effective cloud cover. Aside from being part of a climate mission, one of the objectives of AIRS is to provide sounding information of sufficient accuracy such that when assimilated into a general circulation model, significant improvement in forecast skill would arise. The at-launch algorithm to produce level 2 products (geophysical parameters) using AIRS/AMSU/HSB data, and expected results based on simulation studies, are given in Susskind et al. (2003). The results of that simulation indicate that the sounding goals of AIRS/AMSU/HSB should be achievable. In that simulation, perfect knowledge of the instrumental spectral response functions and the inherent physics of the radiative transfer equations was assumed. Therefore, if the true state of the atmosphere and underlying surface were known perfectly, one could compute the radiances AIRS, AMSU, and HSB would see exactly up to instrumental noise. Susskind et al. (2003) alluded to the fact that this is not the case in reality, and additional terms would have to be included in the retrieval algorithm to account for systematic differences (biases) between observed brightness temperatures and those computed knowing the "true" surface and atmospheric state, and also to account for residual computational errors after that systematic bias is accounted for (computational noise).

In this paper, we show results based on the algorithm we were using to analyze AIRS/AMSU data on November 30, 2004, which we will refer to as Version 4. This algorithm is very similar to the at-launch version, with the major differences attributable to the factors described above, as well as changes in the internally generated quality flags. As in Susskind et al. (2003), the general retrieval methodology, including quality control, is based only on AIRS and AMSU observations, and does not involve use of a colocated GCM forecast model except for its use to provide the surface pressure. The global ECMWF (European Center for Midrange Weather Forecasting) forecast field for a few selected days is used as “truth” to train regression coefficients, which are used once and for all.

HSB failed in February 2003, and HSB radiances are not used in the current retrieval algorithm so as to allow for a continuous climate data record before and after the loss of HSB. Loss of HSB did not appreciably affect the quality of the retrieved data. AMSU A channel 7 was also found to be very noisy and is not included in any of the calculations.

The Goddard DAAC began processing AIRS/AMSU data using Version 4 in April 2005. JPL delivered an earlier version of the algorithm to the Goddard DAAC, Version 3, for the earliest near real time processing of AIRS level 2 products starting in July 2003. In this paper, we outline the differences between the at-launch version and Version 4, and show sample results using Version 4 on data for September 29, 2004, with a particular emphasis on the quality of the retrieved parameters as a function of fractional cloud cover, using a colocated ECMWF 3 hour forecast as “truth”. More details will be published separately in the AIRS Algorithm Theoretical Basis Document (ATBD). Research to further improve the results of analysis of AIRS/AMSU data, leading to a Version 5 algorithm, is continuing.

2. Differences Between the At-Launch Algorithm and Version 4

The differences between Version 4 and the at-launch version of the retrieval algorithm are relatively small. The post-launch channel frequencies were somewhat different from those pre-launch, as expected, as were the channel spectral response functions. Consequently, new Radiative Transfer Algorithm (RTA) coefficients were generated (Strow et al., 2005) to be consistent with the post-launch instrumental conditions. Minor modifications were therefore made to the set of channels used in the retrieval algorithm. The most significant of these resulted from the finding that more channels in the 4.3 μm region were affected by non-local thermodynamic equilibrium (non-LTE) than previously thought. Radiances in these channels are perturbed during the day, and these channels are currently not used in the retrieval algorithm day or night. It was also found that observed channel brightness temperatures for AIRS, as well as AMSU, were biased from those computed using the RTA with the best estimate of the truth. These biases, referred to as “tuning coefficients”, are subtracted from all terms in the retrieval algorithm involving observed minus computed brightness temperatures. New regression coefficients were also generated (Goldberg et al., 2003, Goldberg et al., 2004) based on the relationship between clear column radiances for an ensemble of accepted retrievals on September 6, 2002, January 25, 2003, and June 8, 2003, and colocated 3-hour ECMWF forecast values of skin temperature and temperature, moisture and ozone profiles. The principal component eigenvectors were trained using data from January 15, 2003. A few AIRS channels exhibit a radiometric instability characteristic, known as “popping”, and these channels are excluded from the list of channels used either in the regression or the physical retrieval steps. It was also found that many of the channels used in the at-launch physical retrieval algorithm were not needed in practice, and are no longer used in the physical retrieval steps so as to make the

physical retrieval computationally more efficient with no loss of accuracy. A new concept has also been introduced in terms of quality control, in which different geophysical parameters retrieved from AIRS/AMSU data have different criteria for acceptance.

2.1 Basic Steps in the Retrieval Methodology

The basic steps in the retrieval algorithm are identical to those shown in Susskind et al. (2003), with the exception of Step 15, and are listed below.

1. Use as a starting point the microwave product which agrees with the AMSU A radiances (Rosenkranz, 2000). This is followed by a temperature profile retrieval using AMSU A radiances as well as AIRS radiances for stratospheric sounding channels that never see clouds. As part of this temperature profile retrieval, the surface skin temperature and microwave spectral emissivity is also updated. The geophysical parameters retrieved in this step are called the MW/strat IR retrieval.
2. Determine initial cloud cleared radiances \hat{R}_i^1 using the atmospheric and surface parameters obtained in Step 1, where \hat{R}_i^1 is the derived estimate of the radiance channel i would see if no clouds were present. A cloud parameter retrieval is also performed to help determine which IR channels are not affected by clouds. These cloud parameters are also taken as the final cloud parameters if the combined IR/MW retrieval is not used (see Step 16).
3. Determine the first guess IR surface parameters and temperature-moisture-ozone profile using \hat{R}_i^1 based on a regression step using 1524 AIRS channels (Goldberg et al., 2003, Goldberg, et al., 2004).
4. Produce an improved temperature profile and microwave spectral emissivity starting from the surface and atmospheric parameters determined in Step 3 using the AMSU A channel radiances

and AIRS channel radiances which do not see clouds. The surface skin temperature is not updated as it is estimated better from AIRS radiances than can be determined from AMSU radiances.

5. Determine updated cloud-cleared radiances, \hat{R}_i^2 , taking advantage of the geophysical parameters determined in Step 3. \hat{R}_i^2 is considerably more accurate than \hat{R}_i^1 because the surface and atmospheric parameters obtained from the AIRS regression step are more accurate than those from the microwave first product, especially the infra-red surface spectral properties which are not determined from the microwave retrieval.

6. Perform a surface parameter retrieval using AIRS surface sounding channels along with AMSU channels 1, 2 and 15. This produces a new skin temperature, IR and microwave spectral emissivity, and IR spectral bi-directional reflectance.

7. Determine \hat{R}_i^3 and new cloud parameters using the geophysical parameters determined in Step 6.

8.-11. Use \hat{R}_i^3 to sequentially determine surface parameters, temperature profile, humidity profile, and ozone profile using the appropriate channels. AMSU A temperature sounding channels 3-6 and 8-14 are also included in the determination of the temperature profile. These are called the first pass retrieved products.

12. Update the temperature profile, using only AMSU A radiances and AIRS channel radiances insensitive to clouds. This profile is also used in the application of quality flags and is referred to as the test microwave only retrieval.

13. Using the first pass retrieved products and updated temperature profile, determine \hat{R}_i^4 and the final cloud parameters.

14. Repeat steps 8 and 9 using \hat{R}_i^4 to obtain the final product surface parameters and temperature profile. The initial guess used in the second pass surface parameter and temperature profile retrievals is identical to that of the first pass but all other parameters are updated, such as the clear column radiances, moisture profile, etc. The channel noise covariance matrix is also updated to account for better estimates of the other parameters. In addition, channels in the water vapor band which are highly sensitive to lower tropospheric water vapor are included in the final temperature profile step (but not the first pass) because an accurate moisture profile has now been retrieved. The moisture profile and ozone profile retrieval steps are not repeated, as no appreciable improvement in parameters resulted from further retrieval steps. The geophysical parameters retrieved from this step and the following step are called the combined IR/MW retrieval.

15. Determine the CO profile.

16. Determine whether cloud parameters derived in step 2 (MW/strat IR retrieval) or step 13 (combined IR/MW retrieval) should be reported and used in the computation of OLR. Apply quality flags to all retrieved parameters.

17. Determine OLR and clear sky OLR using the appropriate state, either from Step 13 or Step 2 (OLR is insensitive to the CO profile).

2.2 Channels used in the Physical Retrieval Steps

AIRS contains 2386 spectral channels. Regression coefficients are derived using 1524 channels. Considerably less channels are used in the AIRS physical retrieval steps. There are 25 channels used in the surface property retrieval, 58 channels used in the first pass temperature profile retrieval steps (with an additional 7 water vapor channels used in the second pass temperature profile retrieval step), and 49 water vapor channels, 26 ozone channels, and 20 CO channels are used in the

constituent profile retrieval steps. In addition, 45 channels are used in the cloud clearing step, which are a subset of the temperature profile and surface property channel set in both the longwave and shortwave portions of the AIRS spectrum.

2.3 Other Minor Differences from the At-Launch Version

The basic algorithms for cloud clearing and retrieval of geophysical parameters are identical to those described in Susskind et al. (2003). A few minor details have been modified, primarily with regard to the number of channels used in each retrieval step. In general, less channels are used in the physical retrieval and cloud clearing steps so as to decrease the computational time required to perform the physical retrievals, with little or no change in accuracy. Some modifications have also been made to the damping parameters ΔB_{\max} used in each retrieval step. In addition, a CO profile retrieval step is now included in the retrieval process which is totally analogous to the H₂O and O₃ profile retrieval steps. Details of these modifications are given in the AIRS ATBD.

3.0 Generation of Tuning Coefficients

Steps in the physical retrieval and cloud clearing algorithms involve the difference between observed (cloud cleared) radiances \hat{R}_i and those computed from some geophysical state, R_i^{comp} using the radiative transfer algorithm (RTA) described in Strow et al. (2005). If one had a perfectly calibrated instrument and perfect parameterization of the radiative transfer physics, then, given the true surface and atmospheric state, the observed radiances \hat{R}_i could be calculated up to instrumental noise. Systematic errors in either the calibration of the observed radiances $\hat{R}_{i,\ell}$ (channel i , zenith angle ℓ), or in the computation of radiances $R_{i,\ell}^{\text{comp}}$, would introduce biases in $(\hat{R}_{i,\ell} - R_{i,\ell}^{\text{comp}})$ and

propagate errors into the solution. We attempt to identify these biases and remove their effect by subtracting them from all terms $(\hat{R}_i - R_i^{\text{comp}})$ whenever they occur in the retrieval and cloud clearing processes. This subtraction is done in the brightness temperature domain for both AIRS and AMSU radiances, in a manner analogous to that described in Susskind and Pfaendtner (1989) and used by Susskind et al. (1997) in the analysis of HIRS2 and MSU radiances:

$$\left(\hat{\Theta}_{i,\ell} - \Theta_{i,\ell}^{\text{comp}}\right)' = \left(\hat{\Theta}_{i,\ell} - \Theta_{i,\ell}^{\text{comp}}\right) - \delta\Theta_{i,\ell} \quad 1)$$

where Θ_i is the brightness temperature corresponding to R_i and $\delta\Theta_{i,\ell}$ is the tuning correction.

The tuning corrections are done in the brightness temperature domain because, to first order, a transmittance error will shift the channel weighting function up or down in the atmosphere by a constant altitude (km), resulting in a roughly constant change in computed brightness temperature.

All retrieval and cloud clearing steps involve $\hat{R}_i - R_i^{\text{comp}}$. A small change in radiance is uniquely related to a small change in brightness temperature according to the derivative of the Planck black body function B. Using this relationship, for AIRS channels, the tuned value of $(\hat{R}_{i,\ell} - R_{i,\ell}^{\text{comp}})$ is therefore given by

$$\left(\hat{R}_{i,\ell} - R_{i,\ell}^{\text{comp}}\right)' = \left(\hat{\Theta}_{i,\ell} - \Theta_{i,\ell}^{\text{comp}}\right)' \left(\frac{dB}{dT}\right)_{\hat{\Theta}_{i,\ell}}, \quad 2)$$

and is used in place of $(\hat{R}_{i,\ell} - R_{i,\ell}^{\text{comp}})$ in all retrieval and cloud clearing steps.

The methodology used to generate the tuning coefficients is to identify systematic differences between $\hat{R}_{i,\ell}$ and $R_{i,\ell}^{\text{true}}$, where $R_{i,\ell}^{\text{true}}$ is the radiance computed using the “true” geophysical conditions for channel i and earth location corresponding to zenith angle ℓ . In order to

generate AIRS channel tuning coefficients, cases were selected thought to be unaffected by clouds so as not to have to account for cloud effects on the observed radiances. The 3-hour ECMWF forecast, collocated to the satellite observations, is used as truth, and observations were limited to non-frozen ocean (henceforth referred to as “ocean”), where the value of surface emissivity, needed to compute $R_{i,\ell}^{\text{true}}$, is known reasonably well (Wu and Smith, 1997). A case is classified as non-frozen ocean based on a topography map and the retrieved surface emissivity at 50.3 GHz. Only nighttime cases were selected so as to avoid effects of solar radiation reflected by the surface as well as effects of non-LTE. Ocean night cases on September 6, 2002 were used to determine the biases. These biases had very little scene or zenith angle dependence. Therefore, for AIRS channels, the tuning coefficient for channel i is taken as a constant

$$\delta\Theta_{i,\ell} = A_i. \tag{3)}$$

where A_i is the average bias over all scenes and zenith angles. AIRS tuning coefficients, determined in this manner, are applied only to channels in the CO₂ and N₂O absorption regions (650 cm⁻¹ – 756 cm⁻¹ and 2200 cm⁻¹ – 2420 cm⁻¹). Tuning coefficients determined for the other spectral regions are not used because of uncertainties in the ECMWF water vapor and ozone profiles used as “truth”, as well as uncertainties in the sea surface temperature and spectral emissivity.

Figure 1a shows the tuning coefficients determined for all AIRS channels. They are applied only in those spectral regions indicated by the horizontal bars. Figure 1b shows that daytime biases are very similar to nighttime biases, except for the region between 2240 cm⁻¹ and 2386 cm⁻¹, and greater than 2400 cm⁻¹. Daytime radiances in the first spectral region are affected to varying degrees by non-LTE and in the second spectral region by solar radiation reflected by the surface. Figure 1b indicates by stars the channels currently used for temperature sounding in the spectral

region $2200 \text{ cm}^{-1} - 2420 \text{ cm}^{-1}$, which is a smaller set than in the at-launch version. Channels sensitive to non-LTE effects are not used in the physical retrieval step at this time because non-LTE effects are not currently accounted for in the RTA.

The procedure used to generate AMSU tuning coefficients is analogous to that used to generate AIRS tuning coefficients. The coefficients used were generated for ocean cases on September 6, 2002. These cases were screened to eliminate contamination from precipitating clouds. Unlike the biases found for AIRS channels, AMSU channels had a pronounced, and systematic, zenith angle (beam position) dependence. This arises from effects of antenna sidelobes, which were not adequately accounted for in the calibration of the AMSU observations. Based on this finding, AMSU channels are tuned according to

$$\delta\Theta_{i,\ell} = A_{i,\ell} \tag{4)}$$

with $A_{i,\ell}$ determined using the September 6, 2002 data. The AMSU tuning coefficients are shown in the AIRS ATBD. The AIRS and AMSU tuning coefficients, determined using ocean cases on September 6, 2002, are used globally for all time periods.

An empirical estimate of the uncertainty of the tuning coefficients was also derived in a manner described in detail in the AIRS ATBD. This uncertainty, generally on the order of 0.3K, was estimated by looking at the standard deviation between observed brightness temperatures for clear cases and those computed from the solution using tuned observations. This term is added to the diagonal term of the channel noise covariance matrix used in all retrieval and cloud clearing steps.

4 New Quality Flags

The major change to the at-launch algorithm is a new concept with regard to quality flags. Susskind et al. (2003) discussed a number of threshold tests used to determine whether the combined IR/MW retrieval is of good quality. These tests utilize only the AIRS/AMSU radiance data. No external data, such as GCM forecast fields or MODIS observations are used. If the tests were all passed, the combined IR/MW retrieval state, and associated clear column radiances were reported, as well as cloud and OLR values consistent with the AIRS radiance observations and the IR/MW retrieval state. If any of the tests were not passed, the IR/MW retrieval state was “rejected” and the MW/strat IR retrieval state was reported, as well as associated values of cloud parameters and OLR consistent with that state. Rejection usually implied problems with regard to treating effects of clouds in the field of view, and rejected cases produced generally poorer results in the mid-lower troposphere and at the surface.

The basic approach used now with regard to quality flags is identical with one major exception: different quality flags are used for different geophysical parameters. Problems dealing with clouds in the field of regard (3x3 array of AIRS fields of view) may produce a poor temperature profile in the lower troposphere, but should not degrade accuracy of the stratospheric temperature or upper tropospheric water vapor. For this reason, a less strict threshold test is applied to accept stratospheric temperatures than lower tropospheric temperatures. Cases are classified 0-6 according to their ability to pass six increasingly more stringent threshold tests. The higher the number, the tighter the test which is passed. Class 6 passes the tight sea surface temperature test, Class 5 passes the standard sea surface temperature test, Class 4 passes the lower tropospheric temperature test, Class 3 passes the mid-tropospheric temperature test, Class 2 passes the constituent profile test, Class 1 passes the stratospheric temperature test, and Class 0 fails the stratospheric temperature test.

The final IR/MW retrieval state and associated clear column radiances and cloud and OLR fields are provided for all Classes except for 0, in which case the MW/strat IR state and associated cloud and OLR parameters are reported.

The threshold tests used to assign quality flags are for the most part identical to those in Susskind et al. (2003), with the addition of four new tests. As before, all tests involve only AIRS and AMSU radiances. Susskind et al. (2003) tested for the following quantities: retrieved effective cloud fraction $\alpha\epsilon$ (percent), given by the product of the geometric fractional cloud cover and its cloud emissivity at $11\mu\text{m}$; retrieved total liquid water content W_{liq} (gm/cm^2); the RMS difference between the temperature profile in the lowest 3 km of the atmosphere retrieved in the final state and the test microwave only retrieval, $\Delta T(\text{p})(\text{K})$; the final cloud noise amplification factor, $A^{(4)}$; the final effective cloud noise amplification factor, $A_{\text{eff}}^{(4)}$; the RMS value of the weighted residuals of the clear column brightness temperatures of the channels used in the cloud clearing process, ΔF ; the ratio of the weighted residuals of the channels used to determine the final temperature profile to its theoretical value, R_{temp} ; and the analogous ratio for the channels used to determine the final surface parameters, R_{surf} . Definitions of all quantities referred to above are given in Susskind et al. (2003) and will not be repeated here. Susskind et al. (2003) threshold values for all of these tests are shown in the first column of Table 1. A test is passed if the value of the parameter used in the test is less than or equal to the threshold value. All tests must be passed for the final IR/MW retrieval state to be accepted.

Version 4 threshold values for each of the 6 classes described above for all of these tests are shown in Table 1. Tests for some classes use separate threshold values for ocean cases and land cases. When the thresholds are different, the land threshold is shown in parenthesis, and is always

larger than the ocean threshold or the test is non-applicable. Non-applicable tests are indicated by X. Four new tests have also been added: $A_{\text{eff}}^{(1)}$, which is analogous to $A_{\text{eff}}^{(4)}$ but is applied after the initial cloud clearing; $\Delta\Theta_5(\text{K})$, which is the absolute value of the (tuned) difference between the observed brightness temperature of AMSU channel 5 and that computed from the final retrieval state; $\Delta_{\text{tskin}}(\text{K})$, which is the absolute value of the difference between the regression surface skin temperature and the final surface skin temperature; and RS (Goldberg et al., 2003), which represents how well the observed AIRS radiances can be represented by the use of 200 principle components. Threshold values for these tests for all classes are included in Table 1. Bold values in a class indicate the introduction of a new test or tightening of a previous threshold.

4.1 The stratospheric temperature test

The stratospheric temperature test is the most fundamental test and is used to indicate, first and foremost, whether the final combined IR/MW retrieval, including associated clear column radiances and cloud and OLR parameters should be used, or whether the combined IR/MW retrieval should be “rejected” in all its aspects. The IR/MW retrieval is “rejected” if it is thought to be poorer than the MW/strat IR retrieval, which uses no AIRS channels affected by clouds. The combined IR/MW retrieval cannot always be used because cloud clearing cannot be done under overcast conditions. If the final retrieval were used under such conditions, not only would very poor (too cold) tropospheric and surface skin conditions be derived, but using those conditions to determine cloud fields would result in little or no fractional cloud cover being derived, because AIRS channel radiances computed using the retrieved state would match observed radiances, without the need to add clouds to the scene (see Susskind et al., 2003). Products derived from the combined final

IR/MW retrieval are rejected if the retrieved effective cloud fraction is 90 percent or more. Two tests are added to make sure the clear column radiances are acceptable: $A_{\text{eff}}^{(1)}$ must be less than 200 and RS must be less than 10. Failure of the first test indicates that the initial cloud clearing step had significant problems (note the $A_{\text{eff}}^{(4)}$ threshold was set equal to 8 in Susskind et al. (2003)) and of the second test indicates a significant problem with the observed AIRS radiances (RS equal to 1 is the expected value for nominal radiance performance). Retrieved temperatures 200 mb and above (lower pressures) are flagged as good if this test is passed.

4.2 The constituent profile test

This test is designed to insure that constituent profiles (O_3 , CO , H_2O) are of sufficient accuracy for research use. Constituent profiles are considerably more variable, and less well predicted by models, than are temperature profiles. In general, the more spatial coverage one has, the better, provided the accuracy is adequate. This is especially true with regard to studying interannual variability of monthly mean differences. This applies particularly to water vapor, for which it is desirable to avoid a clear (dry) bias in the selection of the cases to be included in generation of the monthly mean fields. Most CO and H_2O exists in the troposphere, however, and ability to treat cloud effects on the radiances accurately is more important than with regard to stratospheric temperatures. Therefore three tests used in Susskind et al. (2003), designed to indicate potential cloud clearing problems, are included in the constituent profile test as shown in Table 1. The liquid water test threshold is the same as is Susskind et al. (2003), and the thresholds for $A^{(4)}$ and ΔF are considerably less stringent.

4.3 The mid-tropospheric temperature test

Retrieved mid-tropospheric temperatures are affected more by errors in the treatment of clouds in the field of view than are stratospheric temperatures. Therefore, tighter quality control is employed in the mid-tropospheric temperature test. Susskind and Atlas (2004) showed that assimilation of AIRS temperature profiles retrieved from AIRS data, using an earlier version of the AIRS retrieval system (which employed a single rejection threshold for all geophysical parameters), significantly improved forecast skill. Moreover, the improvement was much larger if all accepted cases were used as opposed to use of the slightly more accurate, but much less frequent, temperature soundings in cases found to be clear. Therefore, from the data assimilation perspective, there is a trade-off between accuracy and spatial coverage, as is also true with regard to the study of interannual variability. The thresholds shown in Table 1 are designed to maximize spatial coverage, while minimizing loss in accuracy. Four tests used in Susskind et al. (2003) are now included in the mid-tropospheric temperature test. The first test, $\Delta T(P)$, which contains the difference in the retrieved temperature in the lowest 3 km between the combined IR/MW retrieval and the test MW retrieval, is less stringent than that in Susskind et al. (2003). In addition, the threshold for ΔF has been tightened from that of the constituent profile test, but is still less stringent than in Susskind et al. (2003). Thresholds in the three additional new tests, $A^{(4)}$, R_{temp} , and R_{surf} are all somewhat tighter than in Susskind et al. (2003). Thresholds for $A_{eff}^{(1)}$ and RS have also been tightened from their values in the constituent test, but are still at moderate values. A new test, $\Delta\Theta_5$ has also been added, requiring that the observed brightness temperature for AMSU channel 5, sensitive to lower tropospheric temperatures, should agree with that computed from the combined IR/MW retrieval to

within 2K after tuning is applied. The threshold for ΔF over land is less restrictive than over ocean because ΔF is affected by uncertainty in surface emissivity, which is greater over land than over ocean. R_{surf} and $A_{\text{eff}}^{(1)}$ are also affected significantly by uncertainty in surface emissivity and for this reason, these tests are not utilized over land, so as to maximize spatial coverage. Errors in surface emissivity do not degrade retrieved mid-tropospheric temperatures appreciably. If the mid-tropospheric temperature test is passed, the temperature profile is flagged as good above 3 km of the surface.

4.4 The lower-tropospheric temperature test

Retrieved temperatures in the lowest 3 km of the atmosphere are most sensitive to cloud clearing errors, as well as errors in surface emissivity. A_{eff} and ΔF are both measures of how well cloud clearing is being done and potential problems with surface emissivity. The threshold for ΔF is now tightened considerably and is tighter than in Susskind et al. (2003), in which it had to be relaxed as a compromise so as not to reject the entire profile too often. $A_{\text{eff}}^{(1)}$ is also now used over land, and together with ΔF , flags many cases over arid land (in which retrieved surface emissivity can have large errors) as bad. The test $\Delta_{\text{t}_{\text{skin}}}$ is also introduced which indicates a potential problem with the retrieved surface skin temperature.

4.5 Standard and tight sea surface temperature tests

Sea surface temperature is determined quite well by other instruments such as MODIS. Therefore, for AIRS to produce a useful sea surface temperature product for climate research, it must have very tight quality control. Surface skin temperature is also the product most affected by

errors in the cloud clearing process, especially with regard to very low clouds. In the standard SST Test, thresholds for four tests have been tightened as shown in Table 1. This test is applied only over ocean, as land temperatures are less well measured by other instruments. The test most correlated with sea-surface temperature accuracy was $A_{\text{eff}}^{(1)}$, with lower values indicating more accurate sea-surface temperatures. The percent of accepted sea surface temperatures drops rapidly with decreasing acceptance thresholds however. If $A_{\text{eff}}^{(1)}$ is less than a second threshold, shown for Class 6, then the Tight SST Test is passed.

5. Results Using Version 4

One of the objectives of this paper is to show the extent that high quality soundings and clear column radiances are derived from AIRS/AMSU observations in the presence of clouds. The cloud clearing process does introduce noise in the derived cloud cleared radiances (Susskind et al., 2003). Therefore, one would expect a degradation in retrieval accuracy with increasing cloud cover. It is critical that this degradation should not be appreciable if the retrieved parameters are to be useful for weather and climate research purposes. In this paper, the accuracy of global geophysical parameters derived from AIRS/AMSU observations on September 29, 2004 was evaluated by comparison with a co-located ECMWF 3 hour forecast. Analogous results for September 6, 2002 are shown in Susskind and Atlas (2005). The ECMWF forecast has errors of its own, and this should be borne in mind when interpreting the results of the comparisons. Instead of an assessment of the absolute accuracy of the retrieved quantities, we concentrate on the degree of degradation in “accuracy”, as defined by agreement with ECMWF, occurring with increasing cloud cover. Errors in the ECMWF “truth” may decrease the apparent differences in accuracy between clear and cloudy cases, but only

by making the clear cases appear less accurate than they actually are, and not by making the cloudy cases appear more accurate than they are. In all cases, the quality control methodology described in Section 4.0 is used to include or exclude individual retrieved geophysical parameters from the figures shown.

Figure 2 shows in gray the number of cases for each retrieved effective fractional cloud cover, in 0.5 percent bins, for the whole day September 29, 2004. 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 μm . The average global effective cloudiness was determined to be 44.11 percent. 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 all cases passing the Stratospheric Temperature Test, the Constituent Test, the Mid-Tropospheric Temperature Test, and the Lower Tropospheric Temperature Test, as well as for non-frozen ocean cases passing the standard SST Test and the Tight SST Test. Almost all cases with retrieved effective cloud fraction less than 90 percent pass the Stratospheric Temperature Test, with the percent accepted falling slowly with increasing cloud cover, from close to 100 percent at low cloud fractions to about 65 percent at close to 90 percent effective cloud cover. 79.6 percent of the global cases pass the Stratospheric Temperature Test, with an average effective cloud fraction of 33.08 percent. 78.4 percent of the cases pass the slightly more restrictive Constituent Test, with an average effective cloud fraction of 32.74 percent. 48.5 percent of the global cases pass the Mid-Tropospheric Temperature Test, with an acceptance rate of about 80 percent for low effective cloud fraction, falling to about 20 percent at

80 percent effective cloud fraction, and 10 percent at 90 percent effective cloud fraction. The previous acceptance methodology (Susskind et al., 2003) rejected all cases with effective cloud fraction greater than 80 percent. The mean effective cloud fraction for all cases passing the Mid-Tropospheric Temperature Test is 23.89 percent. Only 26.3 percent of the cases pass the Lower Tropospheric Temperature Test, primarily over ocean, with an acceptance rate near 55 percent for low cloud fractions falling to 5 percent at 80 percent effective cloud fraction and 2 percent at 90 percent effective cloud fraction, and with an average effective cloud fraction of 18.33 percent. The SST acceptance tests are applied only over non-frozen ocean. The standard SST Test accepts 23.3 percent of the ocean cases, and with an acceptance rate of roughly 50 percent under nearly clear conditions, with an average cloud fraction of 9.18 percent, while the Tight SST Test accepts only 10.6 percent of the cases, with an average effective cloud fraction of 5.96 percent. The Tight SST Test allows for more cases than does the clear test (Susskind et al., 2003) which includes only 8.2 percent of the non-frozen cases.

Figure 3a shows the retrieved effective cloud top pressure and effective cloud fraction for ascending orbits on September 29, 2004 in $1^\circ \times 1^\circ$ latitude-longitude bins. The area weighted global mean effective cloud fraction and its spatial standard deviation are indicated in the figure. The results are presented in terms of cloud fraction in 5 groups, 0-20 percent, 20-40 percent, etc. with darker colors indicating greater cloud cover. These groups are shown in each of 7 colors, indicative of cloud top pressure. The reds and purples indicate the highest clouds, and the yellows and oranges the lowest clouds. Cloud fields are retrieved for all cases in which valid AIRS/AMSU observations exist. Gray means no data was observed. Figure 3b shows the retrieved 200 mb temperature field (K). This demonstrates the coverage of cases where stratospheric temperatures are accepted. Gray indicates regions where either no valid observations existed or the stratospheric temperature

retrieval was rejected, generally in regions of cloud cover 90-100 percent. Figure 3c shows retrieved values of total precipitable water vapor (cm). This demonstrates the coverage of constituent profiles. Figure 3d shows retrieved values of 500 mb temperature, demonstrating coverage of accepted mid-tropospheric temperatures. Gaps in the data coverage of mid-tropospheric temperature due to extensive cloud cover are larger than for stratospheric temperatures. Retrieved fields are quite coherent, and show no apparent artifacts due to clouds in the field of view. Water vapor has considerably more fine scale structure than temperature and contains some very large spatial gradients. The extent of gaps in water vapor coverage due to areas of rejected retrievals (retrievals which fail the Constituent Test) are considerably less than with regard to the Mid-Tropospheric Temperature Test, but somewhat larger than with regard to the very loose Stratospheric Temperature Test. As shown in Figure 2, the percent of cases accepted as a function of increasing cloud cover for these two classes of retrievals is almost identical.

Figure 4a shows the difference between the retrieved 700 mb temperature and the ECMWF 3 hour forecast field for ascending orbits on September 29, 2004, for those cases passing the Lower Tropospheric Temperature Test, while Figure 4b shows the same field for all cases passing the looser Mid-Tropospheric Temperature Test. The difference in spatial coverage is significant, particularly over land where 700 mb temperature retrievals appear to be biased warm compared to the ECMWF forecast. Statistics showing the area weighted global mean difference from ECMWF and the spatial standard deviation of the difference are included in the figures. The overall accuracy is somewhat better with the tighter Lower Tropospheric Temperature acceptance criteria, and this difference is significant for data assimilation purposes. When statistics are shown depicting the accuracy of lower tropospheric temperatures (Figures 7 and 8), only cases passing the Lower Tropospheric Temperature Test are included. All data shown in Figure 4b is included in the

generation of lower tropospheric temperature monthly mean fields however, so as to allow for global coverage, especially over arid land regions.

Figures 4c and 4d shows the differences of retrieved ocean surface skin temperature (SST) from the ECMWF SST analysis for the ascending orbits of September 29, 2004. Figure 4c includes only those cases passing the Tight SST Test and Figure 4d also includes those cases passing the standard SST Test. A considerable increase in yield is obtained using the standard SST Test, with some degradation in accuracy of sea surface temperatures. The biases compared to ECMWF are negative in both cases, with a larger negative bias found in cases passing the standard SST Test. Errors due to cloud clearing are typically negative, resulting from under-correcting for effects of clouds in the field of view. This would imply that the Tight SST Test is eliminating more cases where cloud clearing errors are resulting in poorer sea surface temperatures. Caution must be taken however because the ECMWF “truth” may have its own biases.

Figure 5 shows the number of combined daytime and nighttime non-frozen ocean cases between 50°N and 50°S, on September 29, 2004, as a function of the difference of the retrieved SST from the ECMWF analysis in bins of 0.2 K. Results are shown for cases which passed the Tight SST Test, the standard SST Test, and the Lower Tropospheric Temperature Test. Figure 4c and 4d showed the spatial distribution differences for the daytime orbits applying each of the SST Tests. The percent of all non-frozen oceanic cases 50°N-50°S passing each test is included in the statistics, as well as the mean difference from ECMWF, the standard deviation of the difference, and the percentage of outliers, defined as cases passing the test that differ from ECMWF by more than 3K from the mean difference. There is a small negative bias of retrieved Sea Surface Temperatures compared to ECMWF, that increases with increasing acceptance rate, from -0.29K for cases within the Tight SST Test, to - 0.72K for cases passing the Lower Tropospheric Temperature Test. The

standard deviation of the cases from ECMWF also increases slightly. On the other hand, the number of primarily cold outliers increases significantly, from 0.62 percent to 5.90 percent. Therefore, the Lower Tropospheric Temperature Test by itself is not adequate for the purpose of producing accurate monthly mean sea surface temperatures. As with all the test thresholds, experiments are being conducted to optimize the trade-off between spatial coverage and accuracy for best use in studying interannual monthly mean sea-surface temperature differences.

Figure 6a shows RMS differences from the ECMWF 3 hour forecast of retrieved 1 km layer mean tropospheric temperatures, and 3 km layer mean stratospheric temperatures, for non-frozen ocean cases on September 29, 2004. Results shown are for all cases passing the Stratospheric Temperature Test, the Mid-Tropospheric Temperature Test, the Lower Tropospheric Temperature Test, the standard SST Test, the Tight SST Test. Results for those cases passing an additional clear test, as defined by Susskind et al. (2003), are also included in the figure. The number of cases and percentage of all cases included in the statistics are indicated for each test.

Accuracies of retrieved stratospheric temperature, as compared to ECMWF “truth”, improve slightly with increasing stringency of the tests, but are not appreciably different from one another for cases passing any of the quality tests. The large differences from ECMWF above 15 mb are primarily a result of the lower accuracy of the ECMWF “truth” in the upper stratosphere. Tropospheric soundings passing either of the tropospheric quality control tests agree with the ECMWF forecast on the order of 1K. Part of this difference is due to uncertainty in the ECMWF forecast. It is interesting to note that soundings for the 86 percent of the cases for which the Stratospheric Temperature Test was passed are of relatively high quality throughout the troposphere as well, with an RMS difference from ECMWF on the order of 1.7K in the lowest 1 km of the atmosphere. This shows that the cloud clearing methodology works well in up to 90 percent cloud

cover. Nevertheless, the accuracy of all these soundings is not considered high enough for either data assimilation or climate purposes. There is significant further improvement in tropospheric temperature profile accuracy, compared to that for cases passing the tropospheric temperature profile tests, using the smaller subset of cases passing the standard SST Test (23.3 percent of the ocean cases) but relatively little further improvement in those cases passing the Tight SST Test (10.6 percent of the cases), or the additional clear test (8.2 percent of the cases). For data assimilation purposes, we recommend experiments assimilating temperature profiles passing only the standard SST Test, on the one hand, and passing the test for the appropriate for the level of the temperature on the other hand, to assess the trade-off between coverage and accuracy. One might also consider assimilating lower tropospheric temperatures in cases passing the Mid-Tropospheric Temperature Test over ocean to further increase the spatial coverage of the data being assimilated.

Figure 6b shows analogous results for global accepted retrievals, including cases passing the Stratospheric Temperature Test, the Mid and Lower-Tropospheric Temperature Tests, and the clear test (which, over land, ice, and coasts, must also pass the Lower Troposphere Temperature Test). Error statistics in the stratosphere degrade somewhat for cases passing the Stratospheric Temperature Test (79.6 percent of all cases) compared to either of the Tropospheric Temperature Tests (48.5 percent and 25.3 percent). The increase in spatial coverage using the Stratospheric Temperature Test is much more significant globally, compared to using either of the tropospheric tests, than over non-frozen ocean. We therefore recommend using the Stratospheric Temperature Test for stratospheric temperatures for both data assimilation and climate purposes. Global agreement with ECMWF is slightly poorer than over ocean. A much larger difference in agreement with ECMWF occurs between all cases passing the Lower Tropospheric Temperature Test and the Mid-Tropospheric Temperature Test than over ocean, especially in the lower troposphere. For data

assimilation purposes, we feel lower tropospheric temperatures retrieved over land should not be used when the Lower Tropospheric Temperature Test is not passed. Globally, 3.7 percent of the cases passed the clear test, most of which were over non-frozen ocean. Retrievals in these cases are very accurate, but the global spatial coverage is very poor.

Figures 7a and 7b are analogous to Figures 6a and 6b but show statistics only for cases at a given pressure level passing the appropriate quality test. Statistics for cases passing the clear test (identical to those shown in Figures 6a and 6b) are included for comparison. Also shown is the accuracy of the regression first guess temperature profiles for all accepted retrievals and under clear conditions. The accuracy of the physical retrieval is higher than the regression, and more so under cloudy conditions than clear conditions. Part of this is due to the increased accuracy of \hat{R}_i^4 , used to derive the final temperature profile, compared to \hat{R}_i^1 , used to derive the regression first guess.

Figure 7b also includes analogous results determined from the global simulation of AIRS performance shown in Susskind et al. (2003) for all accepted cases (red) and clear cases (pink). In simulation, the truth is known perfectly, while with real data, the 3 hour ECMWF forecast is taken as a proxy for truth. With real data, the degree of degradation for tropospheric accuracy in cloudy retrievals, compared to clear cases, is of the order of a few tenths of a degree, just as it was in simulation. Differences from “truth” are poorer with real data than in simulation however. Two major causes of this degradation are: 1) perfect physics and perfect characterization of the AMSU antenna temperatures were assumed in simulation; and 2) the “truth” has errors of its own in real data. The degradation of sounding accuracy in the presence of “real clouds”, as compared to soundings in clear cases, appears to be similar to that implied by simulation, as does the accuracy of mid tropospheric temperature retrievals for clear cases.

Figure 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. Results are shown for each of the 8 lowest 1 km layers of the atmosphere. Only those cases passing the appropriate temperature profile test 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. This degradation is similar to that shown in an analogous figure in Susskind et al. (2003) for simulated retrievals. RMS temperature differences from ECMWF below 600 mb are somewhat larger than the 1 K goal for retrieval accuracy. Part of this difference can be attributed to the fact that the ECMWF forecast is not perfect. It is also possible that the accuracy of the ECMWF forecast may be somewhat poorer with increasing cloud cover.

Figures 9a and 9b are analogous to Figures 6a and 6b, but show RMS percent difference of retrieved 1 km layer precipitable water from the ECMWF “truth”. In these, and other water vapor statistics, the RMS percent difference weights percent difference in a given case by the “truth”, so as not to inflate percent differences for very dry cases, according to

$$\begin{aligned} \text{RMS percent difference} &= 100 \times \left(\frac{\sum_k (q_k^{\text{tru}})^2 \left(\frac{q_k^{\text{ret}} - q_k^{\text{tru}}}{q_k^{\text{tru}}} \right)^2}{\sum_k (q_k^{\text{tru}})^2} \right)^{1/2} \\ &= 100 \times \left[\frac{\sum_k (q_k^{\text{ret}} - q_k^{\text{tru}})^2}{\sum_k q_k^{\text{tru}^2}} \right]^{1/2} . \end{aligned} \quad 5)$$

where q_k^{ret} and q_k^{tru} are the retrieved and true values of water vapor for case k. These statistics should be used with caution, especially in the mid-upper troposphere, where considerable errors could exist in the ECMWF “truth”. Nevertheless, over ocean, statistics are not appreciably different

for cases passing the different tropospheric and ocean skin temperature thresholds. As with regard to temperature, a larger degradation occurs in agreement of humidity profile with ECMWF in the mid-lower troposphere over land when the looser constituent profile criteria are used. We recommend at this time to use the appropriate temperature test when attempting to assimilate water vapor at a given level of the atmosphere. Soundings passing either tropospheric temperature test also pass the constituent profile test because the temperature profile criteria are equal to, or tighter than, those in the constituent profile test. For climate purposes, we recommend including all cases passing the Constituent Test in the generation of the level 3 product, so as to minimize a dry bias in the sample.

Figures 10a,b are analogous to Figures 7a,b and show water vapor percent differences from “truth” for clear cases and cases passing the temperature test for the appropriate level. Figure 7b includes analogous results found in simulation (Susskind et al., 2003). There is not a significant difference in water vapor retrieval accuracy occurring between clear cases and all cases passing the appropriate temperature profile test with real data, as in simulation.

Figure 11 is analogous to Figure 8, but for percent differences from ECMWF of 1 km layer precipitable water as a function of retrieved effective fractional cloud cover. Only soundings passing the appropriate temperature profile test for a given level of the atmosphere (Mid-Tropospheric Temperature Test or Lower Tropospheric Temperature Test) are included in the statistics, as was done in Figure 10. Agreement with ECMWF degrades slightly with increasing cloud cover primarily in the lowest 2 km of the atmosphere, but not appreciably. Part of this could be due to sampling differences, because the AIRS retrievals determine water vapor in the clear portions of the partially cloudy scene, while the forecast values are for the whole scene.

The fundamental parameter used in the determination of geophysical parameters from AIRS/AMSU data is the clear column radiance \hat{R}_i , which represents the radiance AIRS channel i “would have seen” if no clouds were in the field of view. Geophysical parameters are determined which are consistent with \hat{R}_i . Derived geophysical parameters whose accuracy degrades slowly with increasing cloud cover implies that the accuracy of \hat{R}_i also degrades slowly with increasing cloud cover. \hat{R}_i is an important geophysical parameter derived from AIRS in its own right.

Figure 12a shows the mean value of \hat{R}_i (in brightness temperature units) from 650 cm^{-1} to 756 cm^{-1} for all non-frozen ocean cases $50^\circ\text{N} - 50^\circ\text{S}$ on September 6, 2002 passing the Tight SST Test. The most opaque portion of the spectrum is near 667.5 cm^{-1} , and is primarily sensitive to atmospheric temperatures near 1 mb (50 km). Radiances in the surrounding spectral region are also primarily sensitive only to stratospheric temperatures and are not affected by clouds in the field of view. Radiances at frequencies greater than 690 cm^{-1} see increasing amounts of the troposphere, especially between absorption lines (the locally higher brightness temperatures) and are increasingly affected by cloud cover with increasing frequency. Radiances between lines at frequencies higher than 740 cm^{-1} are also increasingly sensitive to contributions from the ocean surface.

Figures 12b and 12c show the mean and standard deviation of the (tuned) differences between \hat{R}_i and R_i computed from the “truth” for all cases in this geographic domain passing the Tight SST Test, the standard SST Test, the Lower Tropospheric Temperature Test, and the Mid-Tropospheric Temperature Test, respectively. Figure 12c also contains the channel noise spectrum. In this calculation, the “truth” is taken as the ECMWF forecast of temperature-moisture-ozone profile, along with the ECMWF ocean surface skin temperature. The Masuda Ocean surface

emissivity model (1988), revised by Wu and Smith (1997), was used to generate the ocean surface emissivities in the calculation of the expected true radiances, assuming a surface wind speed of 5 m/sec. The surface contribution is the biggest uncertainty in the computation of the “truth” radiances because of errors in both the true ocean skin temperature and in the true surface emissivity.

It is apparent that the difference of clear column radiances from those computed from the truth increases only slightly in the more difficult cloud cases, and in general matches expected radiances to within the AIRS noise level. Standard deviations of observed minus computed brightness temperatures for stratospheric sounding channels are actually lower than the channel noise, because radiances of a AIRS fields of view are averaged together to produce the cloud cleared radiances. The increasing difference of clear column radiances from those computed from the “truth” between absorption lines above 740 cm^{-1} has a large component arising from errors in the “truth”.

It is noteworthy that the biases of observed minus computed brightness temperatures are essentially zero for all cases, with some small negative biases between absorption lines at the frequencies sensitive to the lowest portions of the atmosphere in cases passing the Mid-Tropospheric Temperature Test as a result of small cloud clearing errors in these cases. First of all, this implies that the tuning coefficients derived from clear ocean night ocean cases on September 6, 2002 are equally well applicable to a much larger ensemble of ocean cases on the same day. Secondly, it demonstrates that clear column radiances for cases passing the Mid-Tropospheric Temperature Test are essentially unbiased at most sounding channel frequencies. The standard deviations of the clear column radiances from “truth” are also only slightly dependent on the degree of cloud contamination. Errors in the “truth” dominate the standard deviations shown in Figure 12c,

especially at 667.5 cm^{-1} , which is primarily sensitive to 1 mb temperature, and at frequencies sensitive to the ocean surface. In addition, the larger standard deviation at 679.31 cm^{-1} is a result of significant absorption by O_3 , and those at 729.0 cm^{-1} , 730.8 cm^{-1} , and 745.1 cm^{-1} , and 754.4 cm^{-1} result from significant absorption by H_2O .

Figure 13 shows histograms of the difference between observed and computed brightness temperatures for the two channels indicated by the black dots in Figure 12, at 724.52 cm^{-1} and 749.19 cm^{-1} respectively. These frequencies are primarily sensitive to temperatures at 580 mb and 900 mb respectively, with a large surface contribution at 749.19 cm^{-1} . Results are shown for the four most stringent quality tests. The differences between the accuracy of clear column radiances at 724.52 cm^{-1} , for cases passing the different quality tests with spatial coverage ranging from 9.14 percent to 58.27 percent, are miniscule, with essentially no outliers in any category. Differences are somewhat larger at 749.19 cm^{-1} , but increase only slightly for cases passing the Mid-Tropospheric Temperature Test. For this reason, all clear column radiances are flagged as good for those cases passing the Mid-Tropospheric Temperature Test.

It is apparent from Figure 12 that the tuning coefficients derived for clear ocean night cases on September 6, 2002 are applicable to all ocean night cases on that day. Figures 14a-c show analogous results for all (global) cases passing the Mid-Tropospheric Temperature Test on September 6, 2002 and January 25, 2003 corresponding to a different season and year. The biases (necessary tuning) are shown to be globally applicable, and also constant in time. Standard deviations from the truth at channels more sensitive to the surface are somewhat larger than for the non-frozen ocean cases because of larger errors in the "truth" arising from greater uncertainty in both surface skin temperature and spectral emissivity. The sounding results for September 2004

shown in this paper further demonstrates the stability of the tuning coefficients derived from September 2002 observations.

Operational numerical weather prediction centers currently assimilate radiance observations from IR sounders only for those cases thought to be unaffected by clouds (McNally, et al. 2000). This criterion severely limits the number of IR channel radiances being used in the assimilation processes, and tends to minimize the potential improvement in forecast skill achievable from optimal use of AIRS radiance observations. We encourage operational centers to attempt to use AIRS derived clear column radiances in their assimilation, applying the same quality control so as to accept only those clear column radiances “thought to be unaffected by clouds.”

6. Summary

The AIRS Science Team algorithm used to analyze AIRS/AMSU data to derive geophysical parameters, including cloud cleared radiances, has been described. This methodology is essentially the same as that developed by Susskind et al. (2003) based on experience with simulated data. This algorithm is called Version 4. The major modifications involve treating the effects of errors in the Radiative Transfer Algorithm and a new concept of geophysical parameter dependent quality flags. Accurate soundings of temperature and moisture profiles, and corresponding clear column radiances, are produced in up to 90 percent effective fractional cloud cover. Accuracy of results, as judged by agreement with the ECMWF forecast, degrades only slightly with increasing cloud fraction when appropriate quality control is applied, in manner completely consistent with what was predicted based on simulations. The percent of accepted geophysical parameters does decrease with increasing cloud fraction, in a manner which depends on the geophysical parameter.

The Goddard DAAC began analyzing near real time AIRS/AMSU data, using the Version 4 algorithm described in this paper, on April 1, 2005. The DAAC also began analysis of historical AIRS/AMSU data going back to September 1, 2002, using the Version 4 algorithm. Level 1B (radiances), Level 2 (spot by spot retrievals) and Level 3 (gridded) data is available. The Level 3 data is given on a $1^{\circ} \times 1^{\circ}$ latitude-longitude grid, and averaged in 1 day, 8 day, and monthly mean segments, with ascending (1:30 PM local time) and descending (1:30 AM local time) orbits gridded separately. The DAAC Level 3 data is gridded so as to include retrieved cloud and OLR parameters for all observed cases; temperatures 200 mb and higher in the atmosphere for all cases passing the Stratospheric Temperature Test; water vapor and ozone fields for all cases passing the Constituent Test; temperatures below 200 mb and land surface skin temperature for all cases passing the Mid-Tropospheric Temperature Test; and non-frozen ocean surface skin temperature for all cases passing the Standard SST Test. Examples of 1 day gridded ascending orbits are shown in this paper. The data can be ordered at <http://daac.gsfc.nasa.gov/data/datapool/AIRS/index.html>. Collection 003 should be requested which has results derived using the Version 4 algorithm described in this paper. Collection 002 has results derived using an earlier algorithm (Version 3). These should not be used for scientific studies.

Research is continuing to improve the AIRS/AMSU retrieval algorithm. Version 5 should be made operational and delivered to the Goddard DAAC in 2006. Correcting the current limitations in retrieved land surface spectral emissivity and improving the quality of the error estimates of retrieved geophysical parameters are the highest priorities for the Version 5 algorithm. We also plan to use the improved error estimates directly as quality control indications, rather than using thresholds based on a number of different tests as done in Version 4. Research is continuing to optimize all other aspects of the retrieval algorithm as well.

References

- Aumann, H., M. T. Chahine, C. Gautier, M. Goldberg, E. Kalnay, L. McMillin, H. Revercomb, P. Rosenkranz, W. Smith, D. Staelin, L. Strow, and J. Susskind, AIRS/AMSU/HSB on the Aqua mission: Design, science objectives, data products, and processing systems. *IEEE Trans. Geosci. Remote Sensing*, **41**, 253-264, February 2003.
- Goldberg, M. D., Y. Qu, L. M. McMillin, W. Wolff, 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, February 2003.
- Goldberg, M.D., C.D. Barnet, W. Wolff, L. Zhou, M. Divakarla, Distributed real-time operational products from AIRS. *SPIE International Symposium Optical Science and Technology*, 49th Annual Meeting, Denver, CO, USA, 2004.
- Masuda, K., T. Takashima, and Y. Takayama, Emissivity of pure and sea waters for the model sea surface in the infrared window region. *Remote Sensing of the Environ.*, **24**, 313-329, 1988.
- McNally, A., J. Derber, W. Wu, and B. Katz, The use of TOVS level 1B radiances in the NCEP SSI Analysis system. *Q.J.R. Meteorol. Soc.*, **126**, 689-724, 2000.
- Rosenkranz, P. W., Retrieval of temperature and moisture profiles from AMSU-A and AMSU-B measurements. In *Proc. IGARSS*, 2000.
- Strow, L., S. Hannon, S. De-Souza Machado, and D. Tobin, Validation of the AIRS radiative transfer algorithm. Submitted to *JGR*, 2005.
- Susskind and Pfaendtner, Impact of interactive physical retrievals on NWP. Proceedings of a workshop held at ECMWF, May 9-12, 1989, European Center for Medium-Range Weather Forecasts, Shinfield park, Reading TG2 9AX, United Kingdom, 1989.

- Susskind, J., P. Piraino, L. Rokke, L. Iredell, and A. Mehta, Characteristics of the TOVS Pathfinder Path A dataset. *Bull. Amer. Meteor. Soc.*, **78**, No. 7, 1449-1472, 1997.
- Susskind, J., 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. Geosci. Remote Sensing*, **41**, 390-409, February 2003.
- Susskind, J. and R. Atlas, Atmospheric soundings from AIRS/AMSU/HSB. *Proc. SPIE Conference*, **5425-31**, 311-319, April 12-15, 2004.
- Susskind, J. and R. Atlas, Atmospheric soundings from AIRS/AMSU in partial cloud cover. *Proc. SPIE Conference*, **5806-61**, 587-598, March 28-April 1, 2005.
- Wu, X. and W. L. Smith, Emissivity of rough sea surface for 8-13 μm : Modeling and verification, *Appl. Opt.*, **36**, 2609-2619, 1997.

Figure Captions

Figure 1a. Biases of observed AIRS brightness temperatures minus those computed from the truth for clear ocean night observations on September 6, 2002. Spectral regions where tuning coefficients are applied are given by dark bars.

Figure 1b. Difference of clear ocean dry biases and clear ocean night biases for September 6, 2002. Channels between 2200 cm^{-1} and 2400 cm^{-1} used in the physical temperature retrieval step are indicated in the figure.

Figure 2. Percent accepted retrievals as a function of retrieved effective cloud fraction for different acceptance tests. The average cloud fraction for all cases is indicated, as well as for all accepted cases for each test.

Figure 3. Values and spatial coverage of cloud parameters, 200 mb temperature, total precipitable water, and 500 mb temperature for ascending orbits on September 29, 2004. Gray indicates missing data, or areas where retrievals fail the appropriate quality test.

Figure 4. Differences from ECMWF “truth”, and spatial coverage using different test thresholds, for both 700 mb temperature and sea surface temperature.

Figure 5. Histogram of number of accepted cases as a function of differences of retrieved sea surface temperatures from ECMWF “truth” on September 29, 2004, for different quality tests. Also

indicated are the mean differences, the standard deviation, the percent of the cases passing the test, and the percent of cases differing by more than 3K from the truth.

Figure 6a. RMS differences of retrieved layer mean temperatures from ECMWF “truth” for all ocean 50°N-50°S cases passing different quality tests on September 29, 2004. The percent of cases accepted, relative to all possible cases, is indicated for each test, as well as the total accepted cases.

Figure 6b. As in Figure 6a, but for global cases.

Figure 7a,7b. As in Figures 6a, 6b, but each level includes all cases passing the appropriate temperature profile test. RMS errors of the regression first guess are also included in the figure. The solid gray curves are identical to those in Figures 6a, 6b. Figure 7b also includes analogous results based on the global simulation done by Susskind et al. (2003).

Figure 8. RMS 1 km layer mean temperature differences from the “truth” on September 29, 2004, as a function of retrieved cloud fraction for all cases passing the appropriate temperature quality test.

Figure 9a,b. As in Figures 6a, 6b, but for percent differences of 1 km layer precipitable water from the ECMWF “truth”.

Figure 10. As in Figure 7, but for percent differences in 1 km layer precipitable water.

Figure 11. As in Figure 8, but for percent differences of 1 km layer precipitable water.

Figure 12a. Mean retrieved clear column brightness temperatures for all non-frozen ocean 50°N-50°S cases on September 6, 2002 passing the Mid-Troposphere Temperature Test.

Figure 12b,c. Mean and standard deviation of the difference of retrieved clear column brightness temperatures from those computed from ECMWF “truth” for all cases passing different quality tests. AIRS channel noise is also indicated in Figure 12c.

Figure 13a,b. As in Figure 5 but for difference of retrieved clear column brightness temperatures for two channels indicated in Figure 12a.

Figure 14a-c. As in Figures 12a-c, but for global cases passing the Mid-Tropospheric Temperature Test for September 6, 2002, and January 25, 2003.

Table 1

QUALITY FLAG TEST THRESHOLDS

<u>Susskind et al. (2003)</u>		<u>Version 4</u>					
Test	Acceptable Profile	1) T(p) good 200mb&up	2) q(p) good O ₃ (p) good	3) T(p) good 3km&up	4) T(p) good above surface	5) SST good Standard	6) SST good Tight
$\alpha\varepsilon$	80%	90%	90%	90%	90%	90%	90%
W_{liq}	.03	X	.03	.03	.03	.01	.01
$\Delta T(p)$	1.25	X	X	2.0	2.0	2.0	2.0
$A^{(4)}$	3	X	8.0	2.0	2.0	2.0	2.0
$A_{eff}^{(4)}$	8	X	X	X	15 (X)	8	8
ΔF	1.75	X	8.0	2.0 (6.0)	1.5 (1.5)	1.5	1.5
R_{temp}	1.0	X	X	0.75	0.75	0.75	0.75
R_{surf}	1.0	X	X	0.75 (X)	0.75 (X)	0.75	0.75
$A_{eff}^{(1)}$	X	200	200	30 (X)	30 (30)	9	5
$\Delta\Theta_5$	X	X	X	2.0	2.0	2.0	2.0
Δt_{skin}	X	X	X	X	1.5	1.5	1.5
RS	X	10	10	4	4	1.2	1.2

$\alpha\varepsilon$ is the effective cloud fraction (%)
 W_{liq} is cloud liquid water (g/cm^2)
 $\Delta T(p)$ represents the difference of retrieved lower tropospheric temperatures between MW only and IR/MW retrievals (K)
 $A^{(4)}$ represents the final channel noise amplification factor (unitless)
 $A_{eff}^{(4)}$ represents the final effective channel noise amplification factor (unitless)
 ΔF represents the quality of the initial cloud clearing fit (unitless)
 R_{temp} represents the degree to which the final temperature profile retrieval has converged (unitless)
 R_{surf} represents the degree to which the final surface parameter retrieval has converged (unitless)
 $A_{eff}^{(1)}$ represents the initial effective channel noise amplification factor (unitless)
 $\Delta\Theta_5$ represents the agreement between the observed AMSU channel 5 brightness temperature and that computed from the final solution (K)
 Δt_{skin} represents the difference between the final surface skin temperature and the regression value ((K)
 RS represents the principle component reconstruction score of the observed AIRS radiances (unitless)
Values in parenthesis are for land if different from ocean

Brightness Temperature BIAS
Observed minus Computed (K)

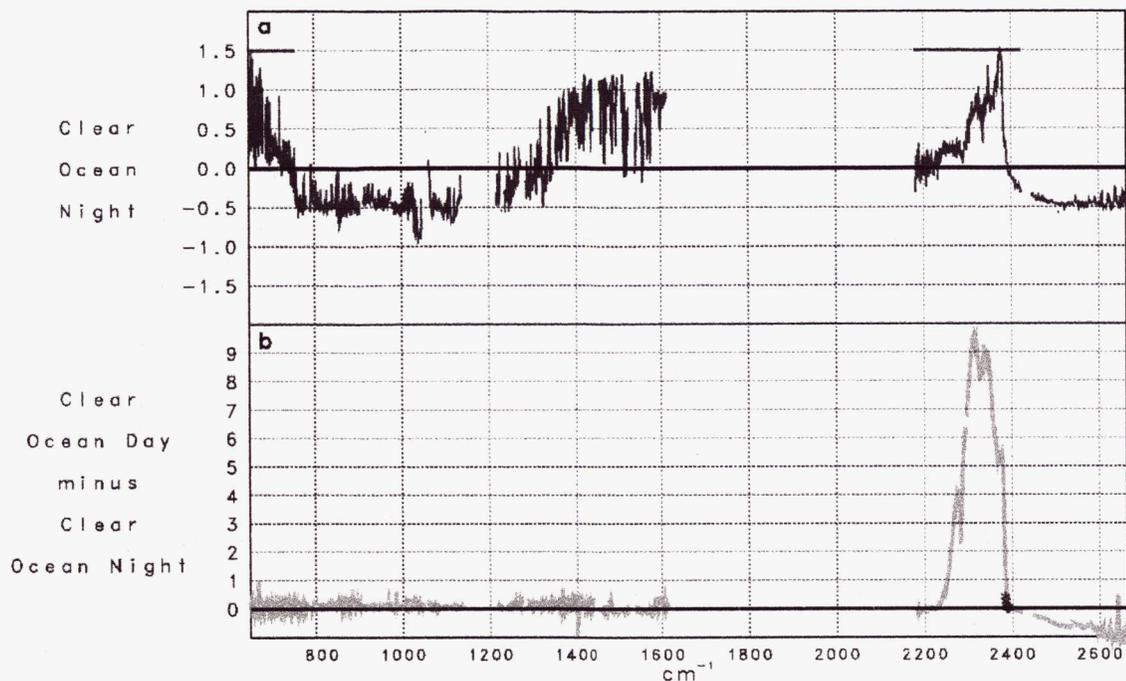


Figure 1

Percent Accepted vs.
Effective Cloud Fraction
September 29, 2004

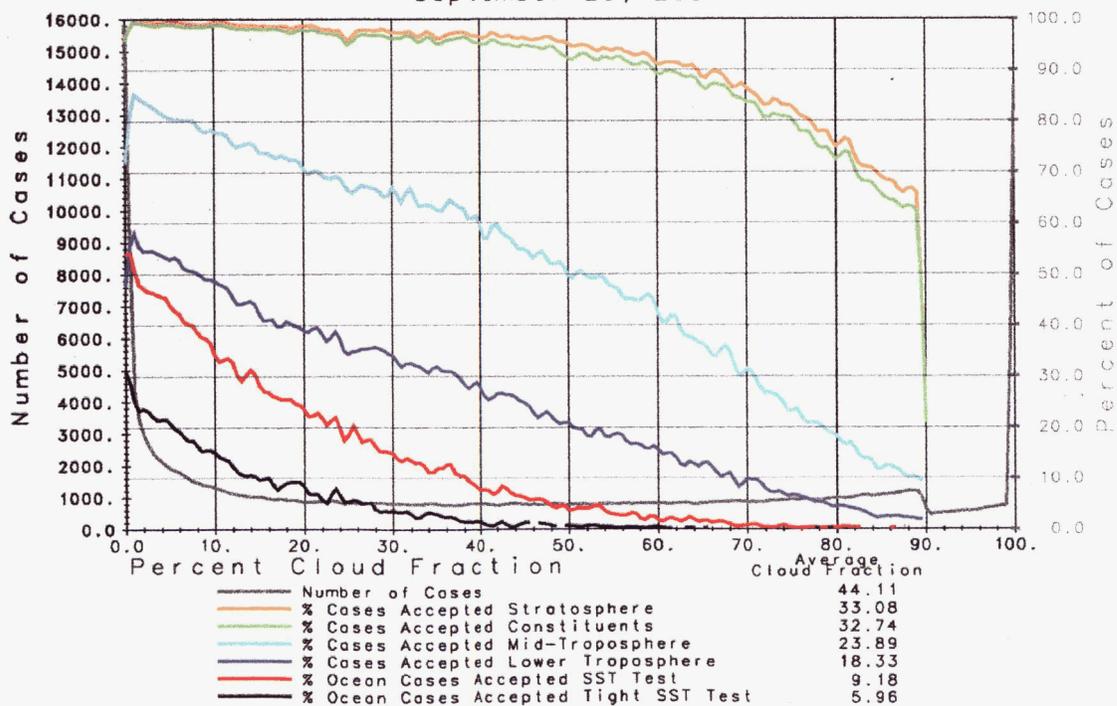


Figure 2

September 29, 2004
1:30 PM

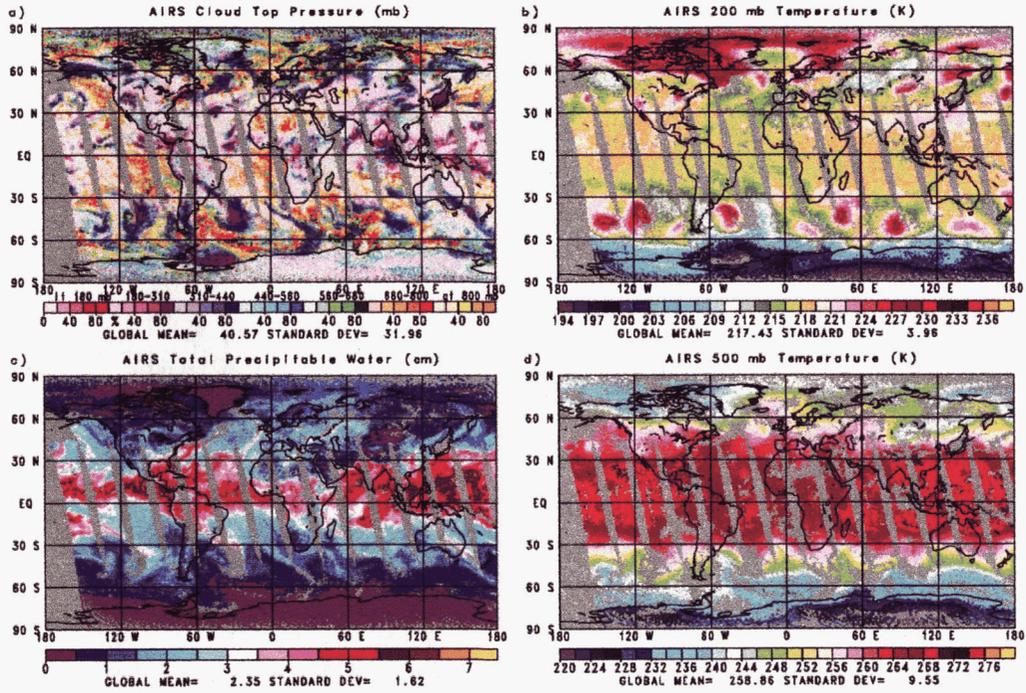


Figure 3

Temperature Differences
AIRS minus ECMWF
September 29, 2004

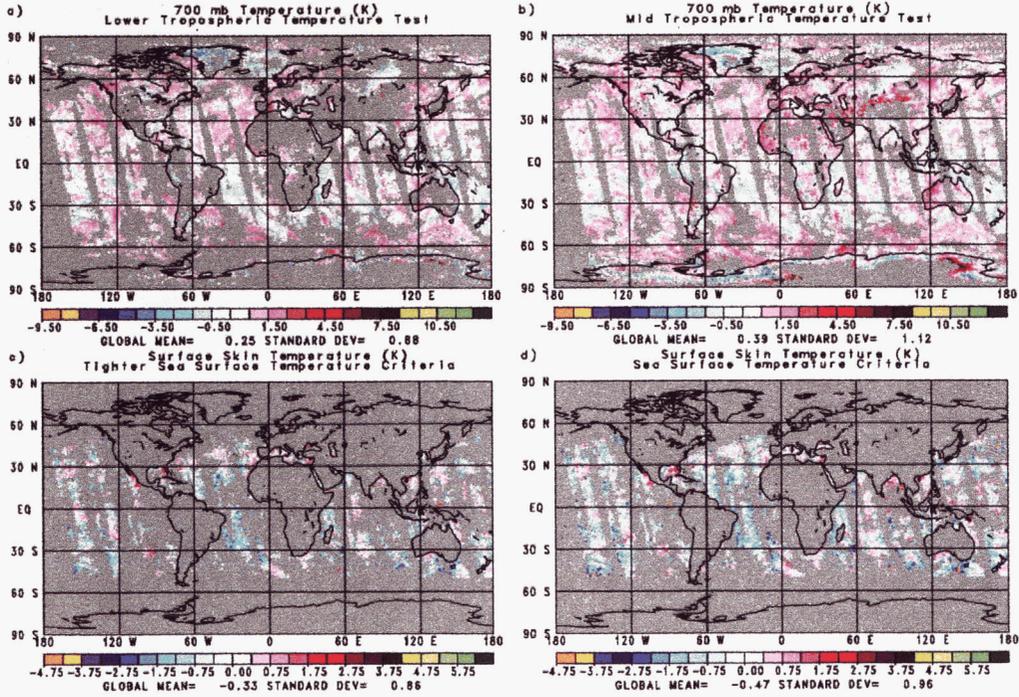


Figure 4

Surface Skin Temperature difference from ECMWF
 September 29, 2004 Daytime and Nighttime combined
 50 N to 50 S Non-Frozen Ocean

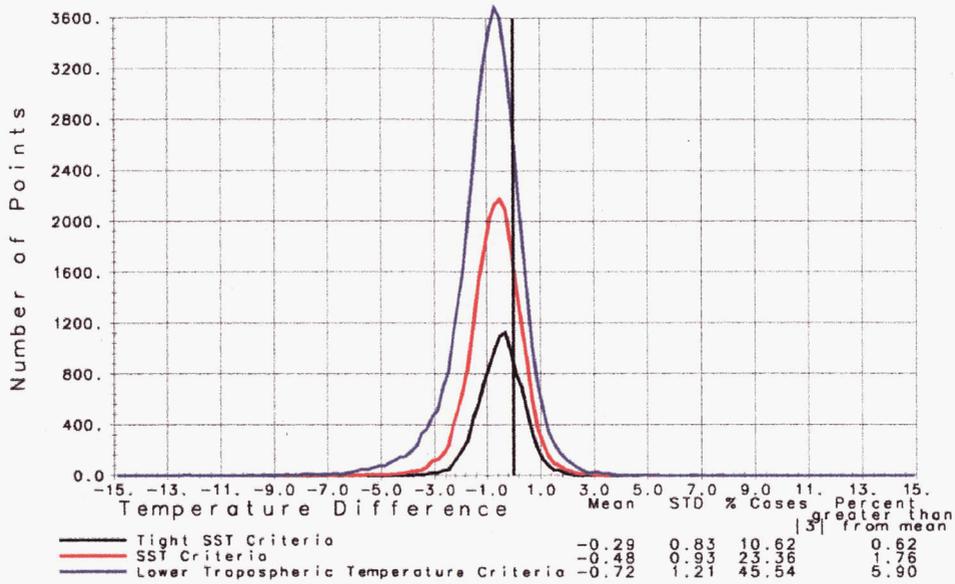


Figure 5

LAYER MEAN RMS TEMPERATURE (°C)
 GLOBAL DIFFERENCES FROM "TRUTH"
 September 29, 2004
 50N to 50S Ocean

LAYER MEAN RMS TEMPERATURE (°C)
 GLOBAL DIFFERENCES FROM "TRUTH"
 September 29, 2004
 Global

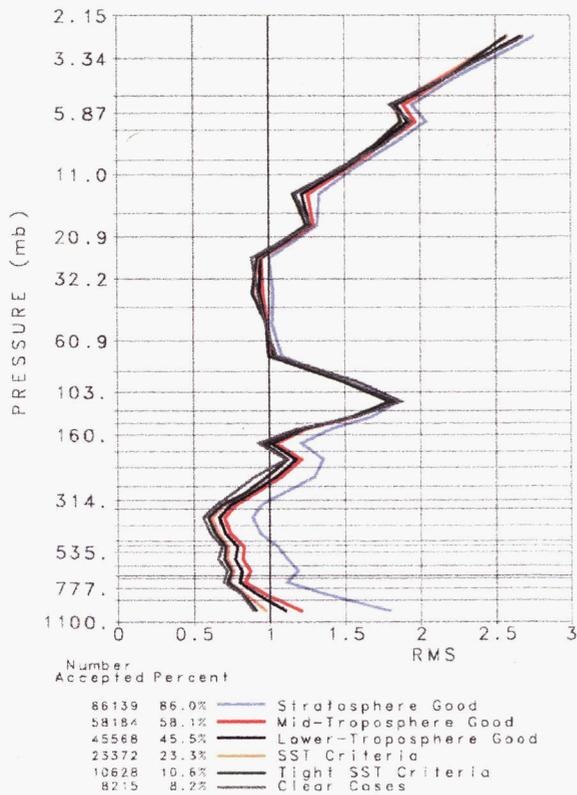


Figure 6a

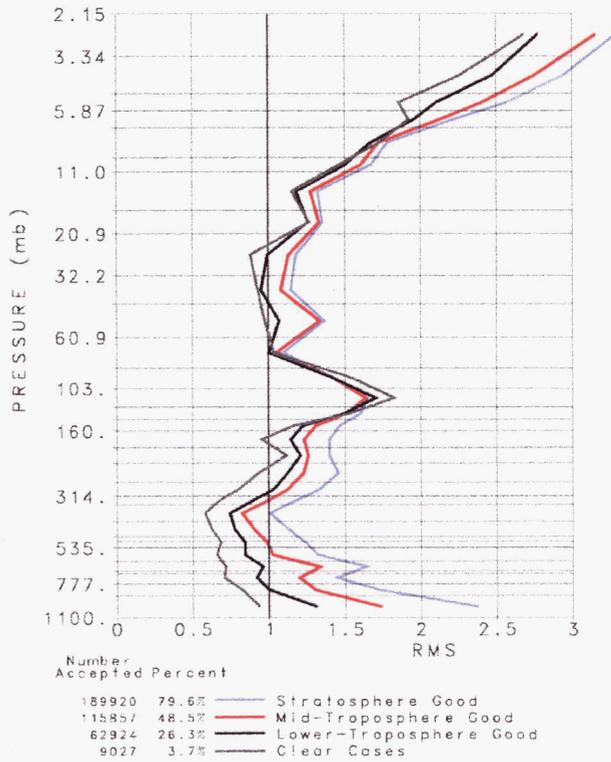
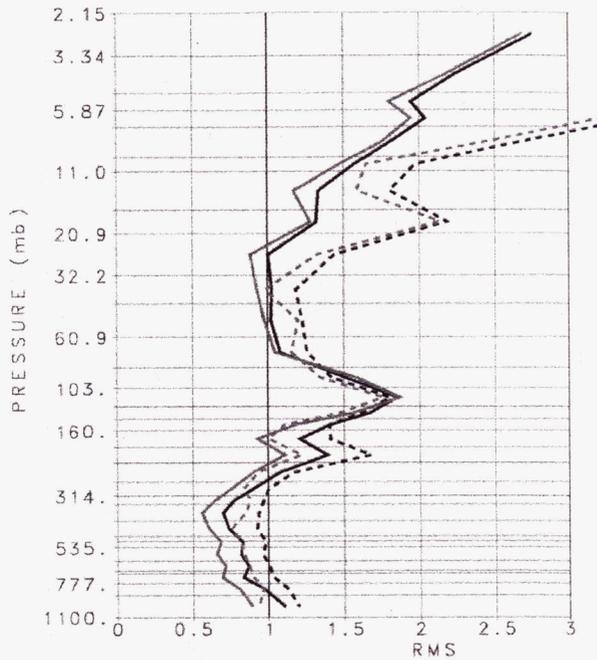


Figure 6b

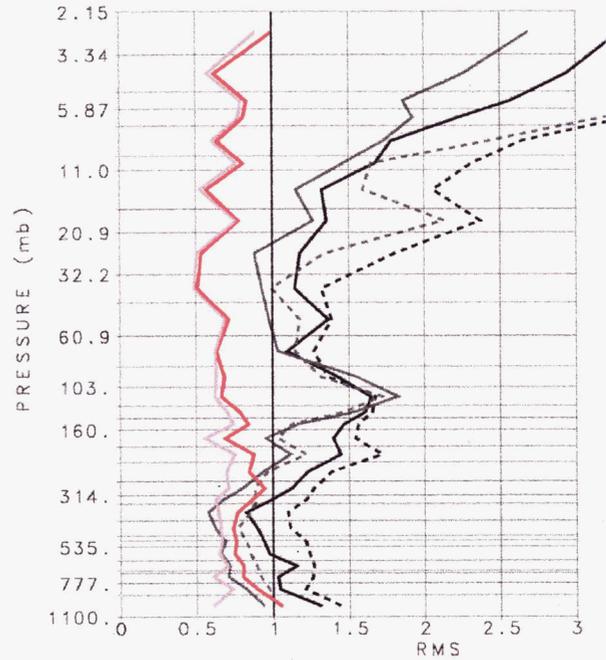
LAYER MEAN RMS TEMPERATURE (°C)
 GLOBAL DIFFERENCES FROM "TRUTH"
 September 29, 2004
 50N to 50S Ocean



Stratosphere	Mid Troposphere	Lower Troposphere	
86.09%	58.15%	45.54%	— All Accepted Cases Final
86.09%	58.15%	45.54%	- - - All Accepted Cases FG
8.21%	8.21%	8.21%	— Clear Cases Final
8.21%	8.21%	8.21%	- - - Clear Cases FG

Figure 7a

LAYER MEAN RMS TEMPERATURE (°C)
 GLOBAL DIFFERENCES FROM "TRUTH"
 September 29, 2004



Stratosphere	Mid Troposphere	Lower Troposphere	
79.64%	48.58%	26.39%	— All Accepted Cases Retrieval
79.64%	48.58%	26.39%	- - - All Accepted Cases FG
3.79%	3.79%	3.79%	— Clear Cases Retrieval
3.79%	3.79%	3.79%	- - - Clear Cases FG
63.90%	63.90%	63.90%	— All Accepted Cases Simulation
6.00%	6.00%	6.00%	- - - Clear Cases Simulation

Figure 7b

AIRS RMS Temperature Difference From Truth
 vs. Effective Cloud Fraction
 September 29, 2005

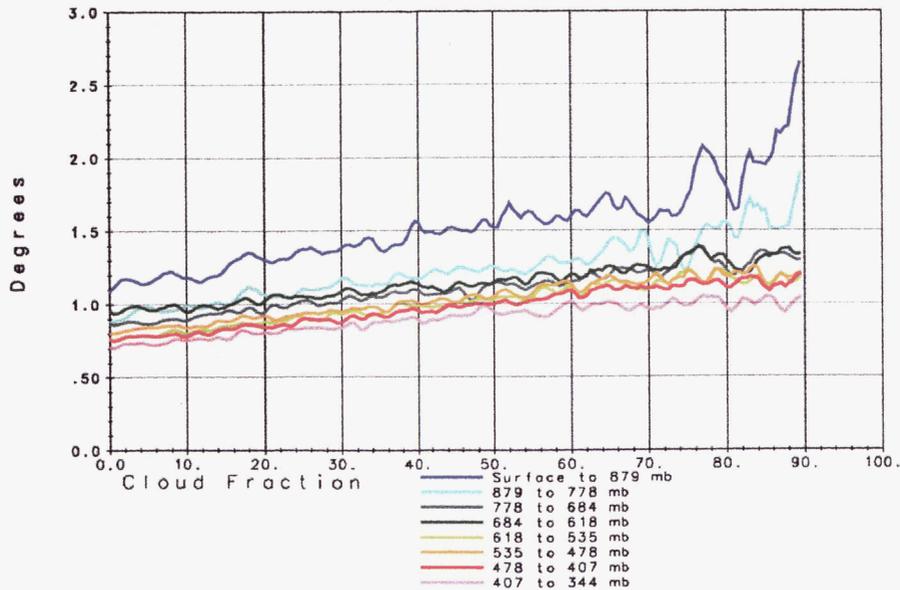


Figure 8

1 Km LAYER PRECIPITABLE WATER
 % DIFFERENCES FROM "TRUTH"
 September 29, 2004
 50N to 50S Ocean

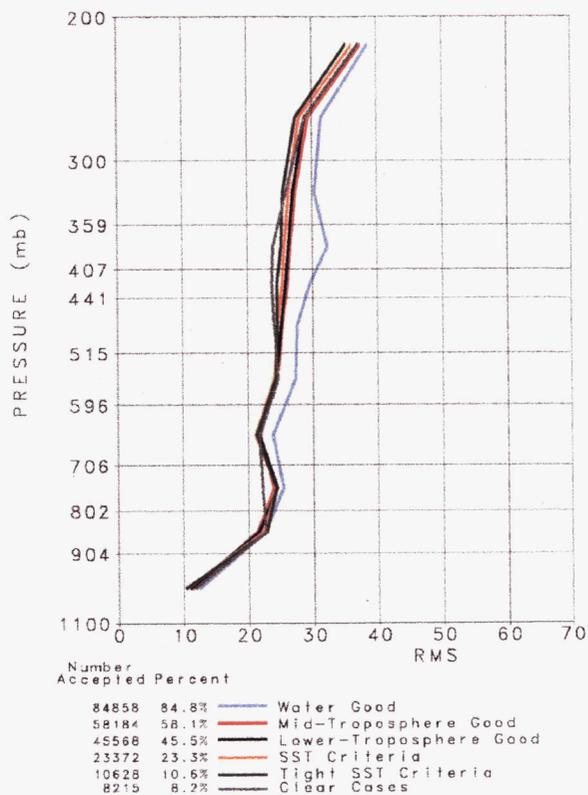


Figure 9a

1 Km LAYER PRECIPITABLE WATER
 % DIFFERENCES FROM "TRUTH"
 September 29, 2004
 Global

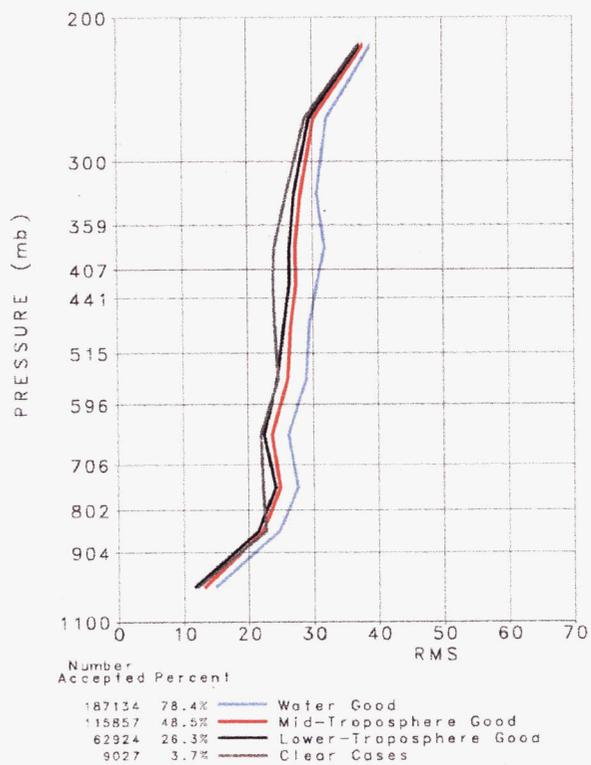
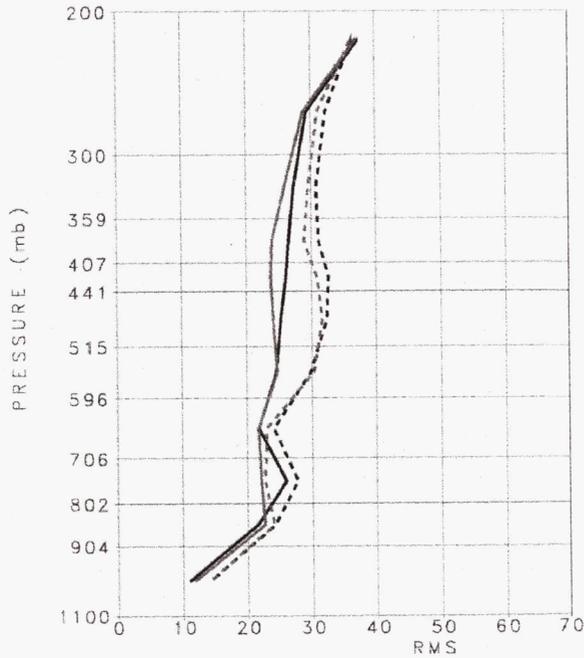


Figure 9b

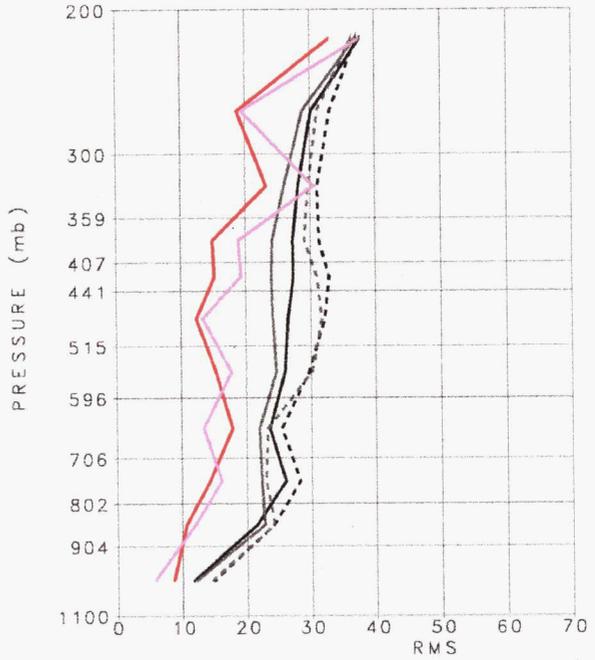
1 Km LAYER PRECIPITABLE WATER
 % DIFFERENCES FROM "TRUTH"
 September 29, 2004
 50N to 50S Ocean



Mid Troposphere	Lower Troposphere	Legend	Category
58.15%	45.54%	—	All Accepted Cases Final
58.15%	45.54%	- - -	All Accepted Cases FG
8.21%	8.21%	—	Clear Cases Final
8.21%	8.21%	- - -	Clear Cases FG

Figure 10a

1 Km LAYER PRECIPITABLE WATER
 GLOBAL % DIFFERENCES FROM "TRUTH"
 September 29, 2004



Mid Troposphere	Lower Troposphere	Legend	Category
48.58%	26.39%	—	All Accepted Cases Retrieved
48.58%	26.39%	- - -	All Accepted Cases FG
3.79%	3.79%	—	Clear Cases Retrieved
3.79%	3.79%	- - -	Clear Cases FG
63.90%	63.90%	—	All Accepted Cases Simulated
6.00%	6.00%	- - -	Clear Cases Simulated

Figure 10b

AIRS RMS Precipitable Water Percent Difference from Truth
vs. Effective Cloud Fraction
September 29, 2004

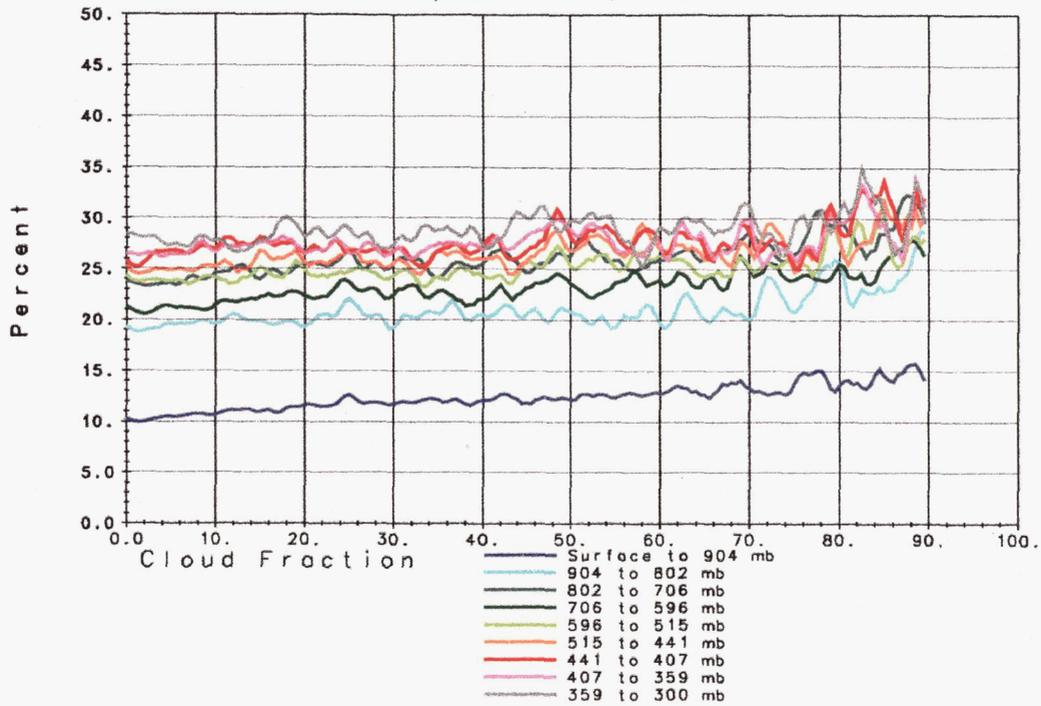


Figure 11

Tuned Clear Column Brightness Temperature minus "Truth"
September 6, 2002 50N to 50S Non-Frozen Ocean
650 to 756 cm^{-1}

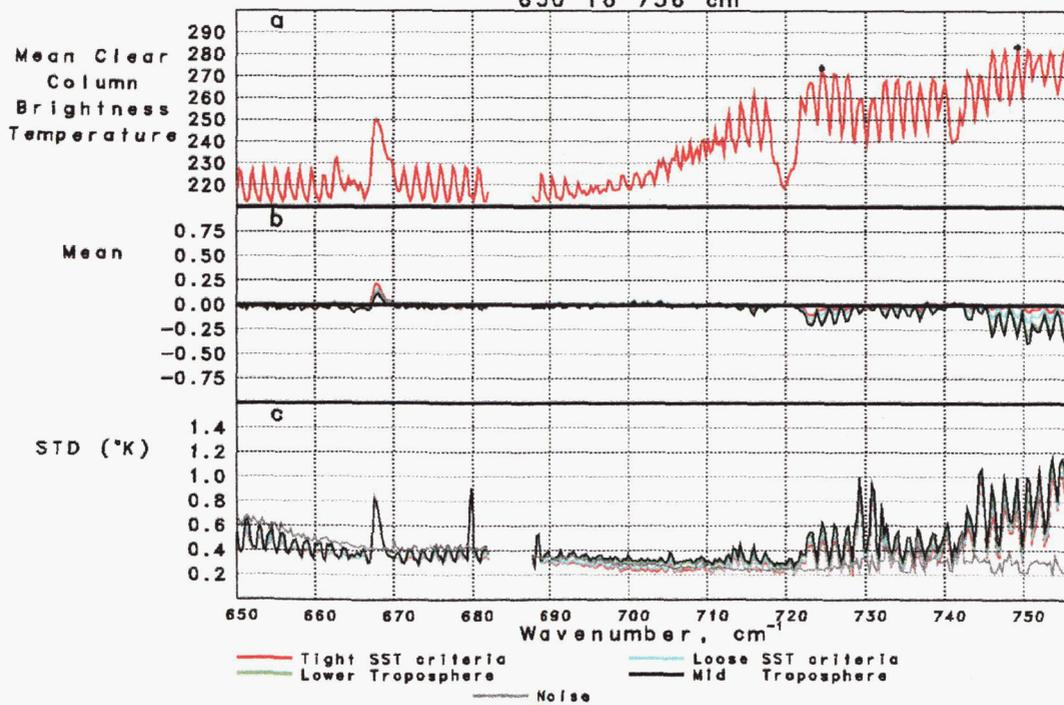


Figure 12

Brightness Temperature difference
 September 6, 2002 Daytime and Nighttime combined
 50 N to 50 S Non-Frozen Ocean
 724.52 cm²

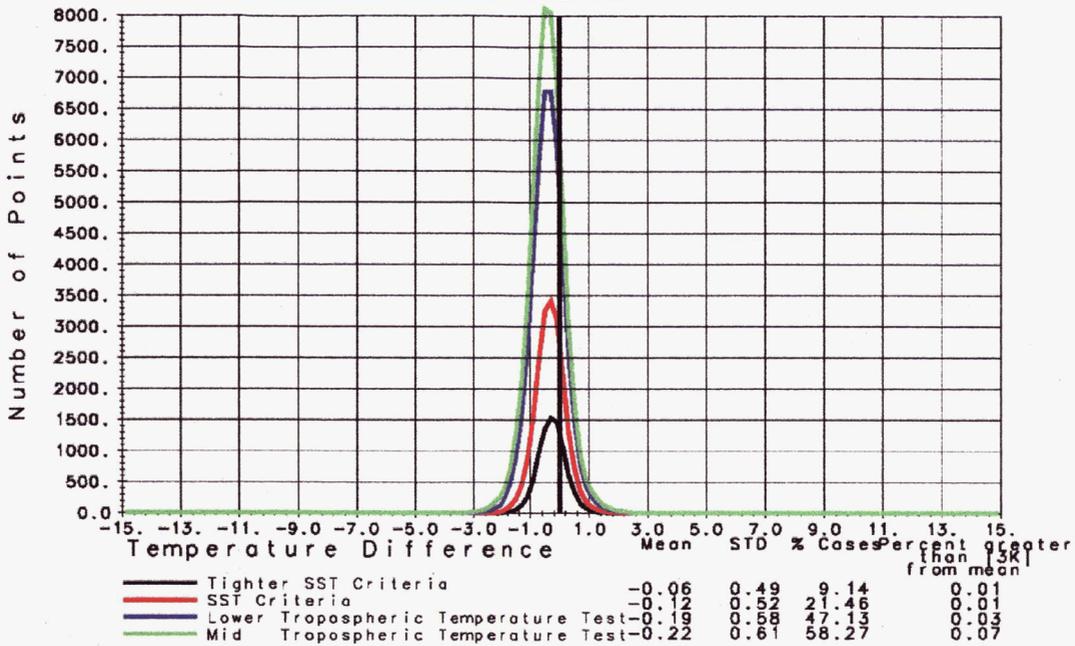


Figure 13a

Brightness Temperature difference
 September 6, 2002 Daytime and Nighttime combined
 50 N to 50 S Non-Frozen Ocean
 749.19 cm²

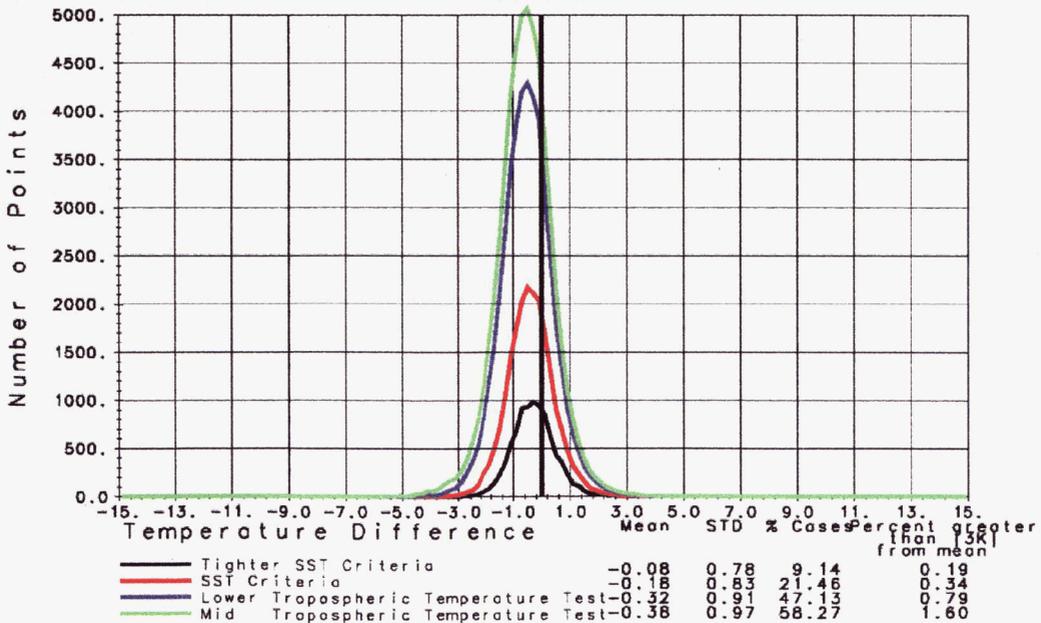


Figure 13b

Tuned Clear Column Brightness Temperature minus "Truth"
Mid-troposphere good criteria Global
650 to 756 cm^{-1}

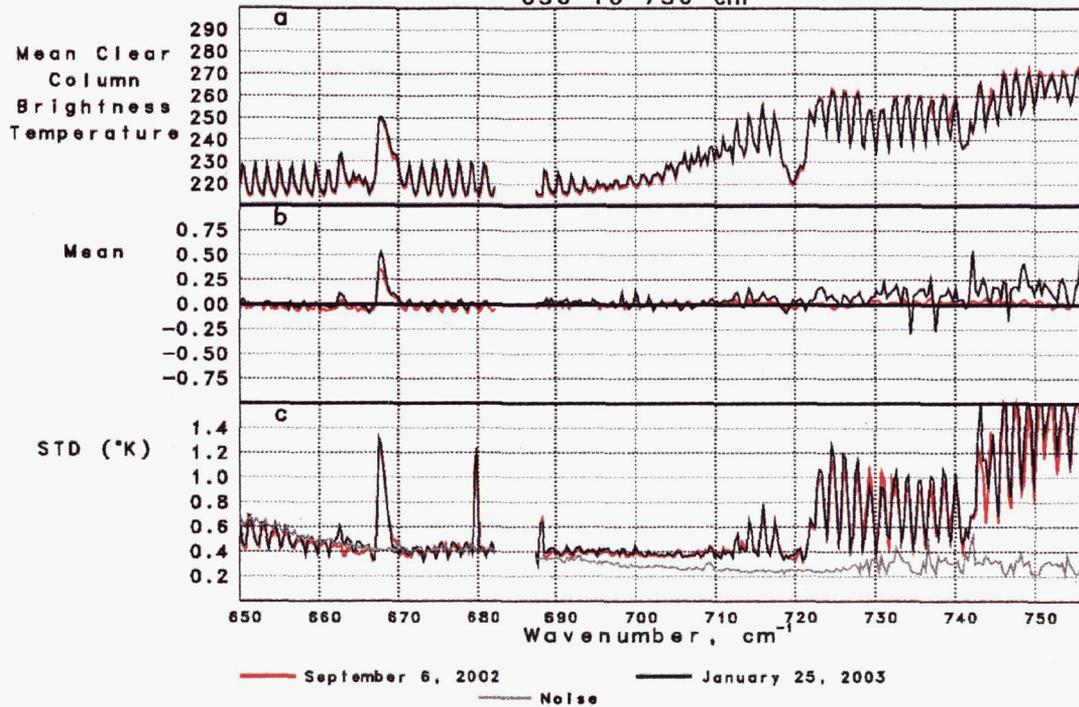


Figure 14