Validation of AIRS v4 ozone profiles in the UTLS using ozonesondes from Lauder, NZ and Boulder, USA

K. P. Monahan,1 L. L. Pan,2 A. J. McDonald,1 G. E. Bodeker,3 J. Wei,4 S. E. George,1 C. D. Barnet,5 and E. Maddy4

Received 24 October 2006; revised 4 June 2007; accepted 15 June 2007; published 13 September 2007.

[1] Ozone sondes observations from Lauder (45.0°S, 169.7°E) and Boulder (39.9°N, 105.3°W) are used to examine the quality of the Atmospheric Infrared Sounder (AIRS) v4 vertical ozone profile product in the upper troposphere lower stratosphere (UTLS). At lower altitudes (~700–200 hPa pressure range), AIRS ozone mixing ratios are larger than ozonesonde measurements, and at higher altitudes (~100–30 hPa pressure range), AIRS ozone mixing ratios are smaller. Compared to the ozonesondes, AIRS retrieval results at Lauder have a median bias of 80% in the region 700–200 hPa, and 0 to –20% in the region 100–30 hPa. For Boulder these values are 40% and 0 to 5%, respectively. Using a tropopause adjusted vertical coordinate system, Lauder has median biases of +90 to +120% in the troposphere and 0 to +25% in the stratosphere whereas Boulder shows median biases of +45 to +70% in the troposphere and 0 to +35% in the stratosphere. Despite the bias, AIRS retrieval in the UTLS region shows a statistically significant positive correlation with the ozonesonde data, indicating that while the absolute values have a large uncertainty, the retrieval captures the variability of ozone in the UTLS region. Hence AIRS ozone is suitable for studies where the change in ozone is important rather than the absolute ozone mixing ratio. Examinations of the training data set show that the retrieval biases are likely influenced by the deficiency of the training data to represent ozone distribution during the regression step of the retrieval. Furthermore the physical retrieval adds little additional information to the final result.


I. Introduction

[2] The upper troposphere lower stratosphere (UTLS) is a region of key interest for understanding global change and chemistry climate coupling. Two-way exchange in the UTLS, known as stratosphere troposphere exchange (STE), is extremely important as water vapor and ozone have large gradients in this region. As water vapor and ozone are radiatively important trace gases, it is essential to have an understanding of the UTLS. Hence the need for ozone measurements is critical. While previous data sets only provide sparse ozone profile data, the high spatial and temporal coverage of the Atmospheric Infrared Sounder (AIRS) has the potential to greatly enhance data availability. This study compares AIRS version 4 (v4) ozone profiles against ozonesonde observations to assess the potential for this product to provide ozone concentration information in the UTLS region. While the ozone product’s main purpose is to improve the quality and accuracy of the temperature product, and hence it is not a focus of the AIRS retrieval, the ozone product may be useful in its own right.

[3] AIRS is located on the EOS Aqua satellite and has been producing twice daily global data sets for a number of different constituents and temperature since late 2002. Previous studies have detailed the AIRS retrieval methodology [Goldberg et al., 2003; Susskind et al., 2003, 2006; Tobin et al., 2006] and have examined the quality of the temperature, water vapor and ozone profile retrievals [Divakarla et al., 2006; Fetzer et al., 2006; Gettelman et al., 2004; Stram et al., 2006; Susskind et al., 2003, 2006]. Fetzer et al. [2003, 2005] review the AIRS validation efforts in detail. In general, validation efforts have focused on the quality of the temperature and water vapor retrievals, and the ozone product has received less attention.

[4] A recent study by Bian et al. [2006] has compared the AIRS ozone profiles with ozonesonde data from Beijing using a new type of ozonesonde called the Global Position-
ing System Ozone sensor (GPSO3). They found that AIRS ozone concentrations appear to be biased high by 20% in the middle troposphere, and biased low by about 20% near the stratospheric ozone peak. Their study also indicated that the agreement between 400–70 hPa, the UTLS region, is mostly within 10%. They also show a consistent variability and positive correlation between the ozonesonde and AIRS in the UTLS region.

[5] The ability to measure ozone variability has widespread application. For example, Randel and Park [2006] have used AIRS to show how transient convective events, associated with the Asian summer monsoon, lead to vertical transport of low ozone and high water vapor into the UTLS region. In Pan et al. [2007], three-dimensional AIRS ozone fields were also effective in providing a large scale background for in situ measurements during a field campaign, where high horizontal resolution AIRS ozone data helped map out the region of large ozone gradient and showed a high degree of correlation with the dynamic features of the UTLS region.

[6] Fetzer et al. [2005] use a limited number of ozonesones at 6 locations to validate the AIRS ozone column and ozone profile. The comparison on AIRS pressure levels showed the stratosphere had about a −10% bias and the troposphere had a bias of between +20% and +70%. The AIRS total column value was within a few percent of the Total Ozone Mapping Spectrometer (TOMS) observations in their study; this is attributed to the positive and negative biases working to cancel one another.

[7] The aim of this paper is to further assess the quality of the AIRS vertical ozone profiles and in particular to examine the potential use of these data for STE studies. It is an extension of previous studies and utilizes two larger and better established data sets.

[8] Section 2 describes the instruments and the data used in this study. A discussion of the comparison methods and related results are presented in section 3. Section 4 describes sets of analyses which test a number of potential explanations for the observed biases between the instruments, including sampling biases and problems with the retrieval technique. Section 5 summarizes the conclusions made.

2. Data Sets

2.1. AIRS Data

[9] AIRS ozone profiles are on 28 layers from 1100 hPa to 0.1 hPa. This study uses the version 4 Level 2 ozone profile data which are available from the Goddard Earth Sciences Distributed Archive Center (http://disc.sci.gsfc.nasa.gov/data/dataset/AIRS/). A detailed description of the AIRS retrieval method is presented in Susskind et al. [2003, 2006]. Briefly, the retrieval method can be separated into three stages: (a) cloud-clearing; (b) regression retrieval; and (c) physical retrieval.

[10] The cloud-clearing stage is necessary as clouds have a significant effect on observed infrared radiances. Therefore an accurate treatment of the effects of clouds on the observed AIRS radiances is critical to obtaining accurate soundings. The cloud clearing processing uses the multiple AIRS observations within the 3 by 3 array of AIRS spots in a single AMSU-A footprint to determine information about the fractional cloud coverage and altitude associated with a number of cloud layers. This allows the radiance that would be observed under clear sky conditions to be obtained (see Susskind et al. [2003] for more details).

[11] The methodology of the regularized regression retrieval (the second step) is described in Goldberg et al. [2003]. This process uses a predetermined relationship between AIRS radiances and atmospheric profiles, including temperature, water vapor and ozone, to produce a first estimate of the profiles. The relationship is a coefficient-matrix obtained using a least squares regression and principle component analysis relating an ensemble of atmospheric profiles and the simulated AIRS radiances. The profile ensemble is created by three days of the European Centre for Medium Range Weather Forecasts (ECMWF) reanalyses (6th September 2002, 25th January 2003, and 8th June 2003) and is referred to as the training data set. The profiles produced by the regression retrieval are then used in the AIRS physical retrieval (the final retrieval step) as the initial profile of iteration. The physical retrieval iterates to minimize the radiance differences between the forward model results and measurements, details are given in Susskind et al. [2003, 2006]. In principal, the training data set can be observationally data based, but a data set that meets the required needs of the training process, i.e., coverage that accommodates different observing geometry of the AIRS instrument, does not exist. One of the goals of this work is to show the potential issues of using ECMWF profiles as a training data set in the regression retrieval. It was recognized during the pre-launch retrieval studies using simulated data, that biases in the training data set can affect the retrieval results, and the physical retrieval step should serve to minimize the biases [Goldberg et al., 2003].

[12] While AIRS makes measurements in 2378 spectral channels, considerably fewer channels are used in the AIRS physical retrieval. Susskind et al. [2006] indicates that 58 channels are used for the temperature profile retrieval, 49 channels for water vapor, and 26 channels for ozone. AIRS data are collected in granules with 1350 profiles per granule. Each granule represents 6 min worth of data. AIRS has a resolution of roughly 50 km by 50 km along the orbital track. Profiles within an AIRS granule are spatially separated on average by 0.15 degrees in latitude and 0.2 degrees in longitude. Hence the AIRS profiles are relatively dense spatially, particularly in comparison to many other data sets. Profiles are selected using the quality and pressure validity flags to ensure that only high quality ozone retrievals are used. In particular, only profiles where the quality flag was equal to zero (i.e. Qual_O3 = 0), indicating the best quality are used. Any measurements at levels that have pressures greater than the valid bottom pressure (i.e. Press_valid_bot) are also removed from the analysis.

2.2. Ozonesonde Data

[13] Ozonesondes are launched from Lauder (45.0°S, 169.7°E) and Boulder (39.9°N, 105.3°W) at approximately weekly intervals. There are 189 Lauder ozonesonde records (available between August 2002 to November 2005) and 111 Boulder ozonesonde records (available between August 2002 to September 2005) that span the AIRS measurement period.

[14] The ozonesondes launched from Lauder are ECC (electrochemical concentration cell) EN-SCI 1Z-series ozo-
nesondes operating with a 0.5% buffered KI cathode solution [Boyd et al., 1998]. Corrections are also applied to the ozonesonde values above 200 hPa to account for pump efficiency degradation [Bodeker et al., 1998]. The response time of ozonesondes results in an overestimation of the altitude by about 150 m [Brinksma et al., 2000]. Due to the low vertical resolution of the AIRS data, the Lauder AIRS/ozonesonde comparisons were found to be relatively insensitive to any altitude adjustment. Hence to maintain consistency with the Boulder ozonesondes, no altitude adjustment was applied.

[15] The ozone profile uncertainty is less than 5% in the troposphere, 2–4% in the stratosphere up to about 30 km, and 4–7.5% between 30 and 35 km Bodeker et al. [1998]. Downwind drift causes negligible differences in ozone measurements.

[16] Boulder ozonesondes are EN-SCI 2Z-type ozonesondes operating with a 2% unbuffered KI cathode solution. They have an accuracy of about 10% for tropospheric mixing ratios, except for very low mixing ratios (i.e., <10 ppbv) when accuracy is 15% Newchurch et al. [2003]. Comparisons of the column ozone from the ozonesonde measurements and those from TOMS overpasses indicate agreement to within 2%.

[17] Ozonesondes and AIRS have two distinct methods for measuring ozone, this results in different measurement ranges for each instrument. On a profile by profile basis, the measurement range also varies due to different factors, such as clouds. The overlapping measurement region between AIRS and the ozonesondes always covers the UTLS, hence this region is the focus for this paper.

3. Comparison of Ozonesonde and AIRS Data

[18] All AIRS profiles to be compared with ozonesondes must be nearly coincident in both space and time to reduce the effect of variations in the ozone field. In this study the following coincidence criteria are applied to select the AIRS profile; The nearest AIRS profile within an approximately 100 km radius of the ozonesonde launch site and the closest available in time are selected. The impact of time and spatial sampling is examined in section 4.1.

[19] The AIRS retrieval provides ozone concentrations on 28 pressure levels. With the exception of the uppermost level, the values reported at the bottom boundary are the mean value between layers [Olsen et al., 2005]. To enable comparison, the high-resolution ozonesonde data are converted to layers on the AIRS grid. There are 6 levels that cover the UTLS region - 250, 200, 150, 100, 70 and 50 hPa.

[20] Ozone mixing-ratio relative difference profiles (d) are calculated using:

\[ d_i = \frac{x_i - r_i}{r_i} \times 100 \]  

(1)

where \( r \) is the reference profile (ozonesonde), \( x \) is the validation profile (AIRS in this case) and \( i \) is the vertical level. This creates a comparison profile for each AIRS/ozonesonde pair.

[21] To provide an overview of the relative differences discussed later in this paper, the mean ozone mixing ratio profile and standard deviation of all of the ozonesonde and AIRS profiles used in the later comparisons are presented in Figure 1. The mean profiles of the measurements from AIRS and the ozonesondes are similar in shape and have similar standard deviations. However, a t-test shows the instrument sample-means to be statistically different on all vertical levels (90–95% significance). It should be noted that the Lauder tropospheric values are smaller than the Boulder tropospheric values.

3.1. Direct Comparison

[22] Comparison of the ozone values on the AIRS pressure levels is the simplest comparison possible. Figure 2 shows relative difference between AIRS and ozonesonde ozone mixing ratios (calculated with equation (1)) on each of the pressure levels that overlap between the instruments. The vertical solid lines give the median bias of AIRS measurements in relation to the ozonesonde measurements. A positive bias indicates that the ozone mixing ratio from AIRS is higher than from the ozonesondes; A negative bias indicates the opposite. In comparison to Figure 1, it may be unexpected that the tropospheric bias should be so large. However, as the tropospheric ozone values are small, the relative differences are extremely sensitive to small changes in ozone, resulting in a large relative bias. Medians and interquartile ranges are used as they are less sensitive to the outliers and skewed distribution of the relative differences observed.

[23] Boulder shows a positive median bias of ~40% between ~700–200 hPa and little bias of 0–5% between 100–30 hPa. This would indicate that for low to mid-stratospheric studies over Boulder, the AIRS data are unbiased. The Lauder comparison shows a large positive median bias of ~80% between 700–250 hPa, and a smaller negative median bias of ~0 to ~20% between 100–30 hPa. The seasonal change in ozone, with regards to the relative difference, is smaller in the lower stratosphere than in the troposphere (figure not shown). Hence the interquartile range in Figure 2 is smaller at lower pressure levels (100–30 hPa) than at higher pressure levels (700–200 hPa).

[24] Fetzer et al. [2005] show a similar range of biases in their comparison. Bian et al. [2006] indicate much smaller biases than this study. However, it should be noted that Bian et al. [2006] use a different type of ozonesonde which displays a 30% bias in the troposphere compared to the conventional ozonesondes used in this study. On adjusting for differences in location and ozonesonde instrumentation, the relative bias in AIRS data observed by Bian et al. [2006] are of comparable magnitude to those seen in the present study.

3.2. Comparison Relative to the Tropopause

[25] While section 3.1 provides some information on the vertical structure of the bias in the AIRS data, it does not account for variations in the UTLS region caused by changes in the tropopause height between profiles, for example, the natural temporal variation in the tropopause height due to the movement of synoptic weather systems. In order to compare only like types of air (e.g., stratospheric air with stratospheric air etc), profiles from the two instruments are aligned using a tropopause relative vertical coordinate system. This improves the stratospheric component of profile comparison and increases the accuracy of the comparison in the UTLS region. The adjustment with respect to
the tropopause acts to remove the seasonal variations in the UTLS region that were present in the pressure level comparison.

[26] Individual tropopause heights are calculated for each temperature profile on geopotential heights from each instrument. To allow continuity between the tropopause calculations, the ozonesonde data are reduced to the same 28 levels as the AIRS data. Tropopause height is calculated at a subgrid scale [Reichler et al., 2003] using the standard thermal definition of the tropopause [World Meteorological Organisation, 1957]. The tropopause level is slightly different for each of the measurements of a AIRS/ozonesonde comparison pair. There are several possible factors that could cause the differences in the tropopause levels calculated, such as slightly different air masses being measured, or differences in the measurement techniques.

[27] Each AIRS/ozonesonde profile pair can then be aligned based on the tropopause level. Any pair where the tropopauses from the two instruments are more than 3 km apart are discarded as either the air masses being sampled are considerably different, or there is large error in the tropopause heights due to the coarse vertical resolution of the AIRS data above 150 hPa. It should be noted that this only removes a small number of comparisons. The tropopause height is set to zero for each profile. This results in the zero point being at a slightly different position with respect to the 28 levels on each of the profiles. In order to easily calculate difference profiles between instruments, it is necessary to have the measurements on the same levels, as opposed to the multitude of levels that result from changing to tropopause relative height. To remove the necessity of interpolating, the ozonesonde profiles on the 28 levels are recalculated from the original high resolution ozonesonde data using the new tropopause relative levels from the corresponding AIRS profile. Relative difference profiles are then calculated using equation (1). Figure 3 displays relative differences using tropopause adjusted heights along with the median and quartile ranges for 2.5 km layers.

[28] This method for aligning tropopauses was selected as it produces the best alignment of the two instruments. If the profiles are incorrectly aligned, for example AIRS stratospheric measurements are not matched with ozonesonde stratospheric measurements, then the resultant profile displays a positive spike below the “common” tropopause level. This was confirmed through Monte Carlo simulations and is a result of the sharp gradients in the tropopause region. There is a slight signal from mismatched tropopauses located at the layer centered at tropopause-2.5 km. This signal is the highest bias value and has a large interquartile range, indicating that there is a large amount of variation in the relative difference values in this region. While the signal causes some reduction in the amount of correlation in this layer, it is of little consequence due to it’s relatively small magnitude.
A comparison of Figures 2 and 3 shows that the same general pattern is observed, however the bias in the troposphere has become slightly larger, and the bias in the lower stratosphere has become smaller and at some levels becomes positive. Lauder has a median bias of +90 to +120% in the troposphere and 0 to +25% in the stratosphere. Boulder has a median bias of +45 to +70% in the troposphere and 0 to +30% in the stratosphere.

The larger bias at the Southern Hemisphere site in comparison to the Northern Hemisphere site shown in Figures 2 and 3 would indicate that there is a factor that is location or hemisphere dependant. A hemispheric instrument bias due to a satellite-specific error is highly unlikely. The quartile range is similar for both locations, this suggests that while the bias is different at the two locations, the variation of the ozone values is similar and displays seasonal patterns. This would indicate that the measurements at each location are consistent apart from the bias.

The correlation between the ozone measurements from AIRS and the ozonesondes is shown in Figure 4 using scatterplots on three relative altitude layers. The correlation coefficients are shown for the three layers used in Figure 4. Over the range of tropopause relative altitudes from tropopause-10 km to tropopause+20 km, the correlation coefficients range from 0.5 to 0.85. The region of maximum correlation is the layer between the tropopause and 10 km above the tropopause, with decreasing values either side of this layer. Hence the results are clearly correlated but they do not lie on the one-to-one line. The shapes of the profiles from the different instruments will be similar but the numerical values will be different. This was shown earlier with regards to Figure 1. This implies that while AIRS does

![Figure 2](image_url)  
**Figure 2.** AIRS/ozonesonde comparison for closest profile averaged on AIRS pressure layers. Black vertical solid lines show the median value and the layer thickness. Horizontal bars display the interquartile range. Crosses show the individual data points and the grey vertical line highlights the zero bias.

![Figure 3](image_url)  
**Figure 3.** AIRS/ozonesonde relative difference comparison for Lauder and Boulder with the altitude adjusted to make the tropopause height the vertical zero baseline. Vertical solid lines show the median value of the differences and the layer thickness. Horizontal bars display the interquartile range. The dashed lines highlight the zero bias in the vertical and the tropopause level in the horizontal. Crosses indicate the individual relative differences. Note a few of the individual differences have been cut off by the horizontal scale selection.
not measure the absolute value correctly, it does accurately represent changes in ozone.

4. Bias Analysis

Figures 2 and 3 show that there is some bias present for the ozone values at both locations, with a larger bias at Lauder. The observations also suggest that the biases observed are generally larger in the lower troposphere than in the stratosphere and that the biases are smallest in the UTLS. The two most likely potential reasons for the patterns observed are either sampling differences at the two sites which are discussed in section 4.1; or biases introduced by the retrieval which could come from any of the three steps detailed in section 2.1. The possible bias from the cloud clearing step is discussed in section 4.2 and from the regression and physical retrieval steps in section 4.3.

4.1. Time and Space Sampling

Multiple AIRS profiles are often within 100 km of the launch site allowing a measure of the geographical variation in the AIRS measurements. The original AIRS profile selected for comparison using the coincidence criteria outlined in section 3 is used as the base profile. All other profiles within 100 km of the launch site are then compared to the base profile using the following formula:

$$base/other\ comparison = \frac{base - other}{\frac{1}{2}\left(base + other\right)} \times 100$$  \hspace{1cm} (2)

Hence each of the original base profiles there are a number of base/other comparison profiles. The mean of all the comparison profiles is zero as the variations in the profiles have a similar magnitude in either direction. The standard deviation of the base/other comparisons gives the best measure of the variation between profiles. The standard deviation is calculated twice, with any extreme outliers (more than three standard deviations) removed from the data set after the first calculation. As the standard deviation shows very little altitudinal dependence, the mean of the standard deviation gives a measure of the bias that could potentially be explained by geographical variations. For Lauder this value is 6.4% and for Boulder is 9.7%.

As indicated previously the Aqua satellite’s orbit provides a roughly twice daily global coverage. From this it follows that no ozonesonde launch should be further than...
6 h from an AIRS measurement, however as a result of the orbital pattern, this is not always true. The Lauder AIRS and ozonesonde profiles are separated by a mean of 4.4 h with a standard deviation of 1.7 h. In comparison Boulder has a mean of 1.9 h with a standard deviation of 1.2 h. The relative difference calculated using equation (1) is compared, on relative altitude layers, to the time separation between the AIRS overpass and ozonesonde launch. An example of the relationship between the time separation and relative difference is shown in Figure 5 for the relative altitude layer from 3.75 km below the tropopause to 1.25 km below the tropopause (tropopause-3.75 km to tropopause-1.25 km). The equivalent figure for each layer displays a similar relationship between time separation and relative difference. A larger time separation should produce larger relative differences as the air mass being sampled by the two instruments is likely to have varied by a larger amount. However, for the time separations used in this study, there is no noticeable change in the relative difference in any relative altitude layer.

4.2. Cloud Clearing

[36] As previously indicated the cloud clearing stage of the retrieval takes the observed radiances and adjusts them to provide estimates of the radiances if no clouds were in the field of view. A useful way to examine whether this stage of the retrieval scheme explains the form of the bias is to examine the relationship between the differences observed as a function of cloud fraction [Suskind et al., 2006]. If the bias is dependent on the cloud fraction, then it would imply that the bias is related to this step. However, comparison on several altitude levels both in the troposphere and the stratosphere shows no relation between the cloud fraction and the bias. From this we can conclude that the bias is not likely to be a result of the cloud clearing stage in the retrieval.

4.3. Examination of the Regression Retrieval

[37] The physical retrieval is the third step of the retrieval process and creates the final profile product. To understand how much the physical retrieval contributes to the final profile product in version 4, we have examined a large number of cases and performed some basic statistical analyses. Figure 6 gives an example using a single profile retrieval. The case selected is for a Lauder ozonesonde, launched 20UT, 24 March 2005. The AIRS retrieval process for this case is shown by three profiles: a reference profile, a first estimation profile associated with the regression retrieval and the final retrieval profile. The reference profile is the mean state of ECMWF based training data set. The first
estimation is produced by the regression step and further used as the first guess for the physical retrieval. It is...different horizontal scales in Figures a and b.

Figure 6. An example of AIRS ozone profile retrieval. The “true” ozone profile (dotted curve) is represented by the ozonesonde measurement, launched from Lauder at 20 UT, 24 March 2005. The AIRS retrieval for this case is shown by 3 profiles: the mean state of the ECMWF training data (dashed curve), result of the regression retrieval (solid black curve), and the result of the physical retrieval (thick solid grey curve).

The amount of extra information diminishes toward the surface.

[39] This indicates that the source of the bias results from stages prior to the physical retrieval, as the physical retrieval has very little impact on the final profile. The physical retrieval has been unable to adjust the profile away from the regression step of the retrieval. It should be noted that the result of the regression retrieval has the correct shape when compared to the ozonesonde profile.

[40] The regression step is based on ECMWF data which is known to have biases [Dethof and Holm, 2004; Morcrette, 2003], therefore the source of the bias may be related to this step. The comparisons shown in section 3 show a much

Figure 7. An ensemble estimation of retrieval information from the regression and the physical retrieval. a) shows the mean and standard deviation, represented by solid curve and error bars, respectively, of the relative change (in percent) of the climatology of the training data set and the ozone profile produced by regression. b) shows the mean and standard deviation of the relative change between the ozone profile produced by the regression and the ozone profile produced by the physical retrieval step. The error bars mark the pressure levels of AIRS ozone profile standard product in the UTLS region. Note the different horizontal scales in Figures a and b.
larger bias at the Lauder site than at the Boulder site. If the ECMWF reanalyses data used to create the regression retrieval profile differs significantly from the local ozonesonde climatology, it may lead to consistent inaccuracies in the AIRS retrieval. This bias is also likely to be location dependant. To examine how representative the ECMWF reanalysis training days are for the situation at Lauder and Boulder, comparisons of the data sets are shown in Figure 8. This figure shows a climatology of the Lauder and Boulder ozonesondes compared with the three training days from the ECMWF reanalyses as discussed in section 2.1. The ECMWF ozone has a positive bias at pressures between ~100 to 800 hPa and a smaller negative bias above 100 hPa. This shows that the ECMWF reanalysis is not a perfect representation of the conditions at Lauder and Boulder, particularly at low altitudes. Figure 8 also displays the mean profiles from the data in Figure 2 for comparison purposes. The comparisons of the ozonesondes to ECMWF and to AIRS are extremely similar. This indicates that the AIRS data are highly dependent on the ECMWF training data as might be expected from the data displayed in Figures 6 and 7.

As the AIRS data has been seen to be strongly affected by the ECMWF training set, the distribution of ECMWF values could provide further insight into the retrieved AIRS values. The ECMWF data were converted from pressure levels to altitude, the tropopause calculated using the thermal definition, then the altitudes were adjusted to be tropopause relative as used earlier in the study. The probability distribution of the ozone mixing ratio for three different layers is shown in Figure 9. Each panel shows the probability distribution of ozone mixing ratio values for Lauder, Boulder and the ECMWF training days. A higher value for the occurrence frequency (vertical axis) indicates the corresponding mixing ratio is more likely to occur than a lower occurrence frequency. It can be seen that the distribution patterns of ozone mixing ratio varies between Lauder, Boulder climatologies and the ECMWF training days. The probability distribution of the ozone mixing ratio for Boulder is more similar to the ECMWF climatology, which indicates that the ECMWF data are a better match for Boulder than for Lauder. In the tropospheric layer (tropical-6.25 km to tropopause-1.25 km) the occurrence frequencies are quite different, with the ozonesondes having a lower ozone mixing ratio than the ECMWF reanalyses. In particular, the Lauder distribution is significantly different from the ECMWF distribution in the troposphere. The comparison of the occurrence frequencies is similar in the tropopause region (tropical-1.25 km to tropopause+1.25 km). The overlap between the distributions with increasing altitude seems to explain the increasing correlation between the AIRS and ozonesonde data set. An additional contribution to the higher bias at Lauder could be the fact that the ozone in the UT is much lower in general compared to NH site (Boulder). The maximum occurrence is ~0.04 ppmv, as shown in Figure 9. It is possible that this value is below the detection limit of AIRS instrument, currently estimated to be 0.06–0.08 ppmv. Results presented in this study should lead to a better characterization of the AIRS ozone detection limit and other measurement sensitivity issues.

Figure 8 shows that the bias between the AIRS and ozonesonde data is similar to that between the ozonesondes and ECMWF. Hence the AIRS data are strongly affected by the ECMWF data. Figure 9 more closely pinpoints that the bias is a result of the distribution of the ozonesondes being significantly different from the ECMWF climatology. These results along with Figures 6 and 7 indicate that in the operational retrieval, the biases created in the regression step are not effectively corrected for in the physical retrieval. Since only 26 channels (out of 2400) are used in the operational retrieval for ozone profile, this could be due to the lack of spectral information in the upper troposphere region in the physical retrieval or that the physical retrieval is over constrained.

Previous comparison studies of the ECMWF ozone product (both the forecast and the ERA-40 models) and ozonesonde data display an overestimation of ozone in the lower troposphere [Dethof and Holm, 2004; Morcrette, 2003]. This problem is smaller during the summer months and at low and middle latitudes. During the winter, the comparison is worse as multiple ozone peaks and reduced ozone levels cause the model to severely overestimate and mislocate ozone. These seasonal variations may explain the seasonal variation in the bias of AIRS/ozonesonde (not shown), where the summer shows a smaller tropospheric bias and the winter has a larger bias.

5. Conclusion

Comparison of the AIRS v4 vertical ozone profiles in the UTLS with ozonesonde observations at Lauder and Boulder has shown the accuracy and sources of error in the data set. Two methods of comparison between AIRS and ozonesondes are displayed. Comparison on pressure levels show that Lauder has a median bias of 80% at low altitudes (~700–200 hPa) and a median bias of 0 to -20% at high altitudes (~100–30 hPa). Boulder has a median bias of 40% at low altitudes (~700–200 hPa) and a median bias of
0 to 5% at high altitudes (~100–30 hPa), as shown in Figure 2. When the vertical coordinate used is a tropopause relative measure, conclusions can be drawn with respect to the troposphere and stratosphere removing the effect of variations in the tropopause level. The tropospheric measurements have a large bias, with a median of +90 to +120% for Lauder and +45 to +70% for Boulder. The stratospheric measurements have a smaller bias, with a median of 0 to +25% at Lauder, and 0 to +35% for Boulder, as shown in Figure 3.

The observations from the Southern Hemisphere site (Lauder) display a much larger troposphere/UTLS bias than those from the Northern Hemisphere site (Boulder). The source of the bias and the cause of this difference were investigated. Uncertainties associated with geophysical variation were found to be 6.4% for Lauder and 9.7% for Boulder. This would only account for small scale variation. The difference in time separation between the instruments is also shown to have a negligible effect. Analysis shows that the bias is likely to be the result of deficiencies in the regression stage of the retrieval scheme. This is exemplified by the difference between the ECMWF training days probability distribution and the ozonesonde probability distribution almost matching the bias in the AIRS/ozonesonde comparison (see Figures 8 and 9). The physical retrieval adheres very closely to the result from the regression retrieval in the upper troposphere (see Figure 7b).

While significant biases exist in the AIRS ozone data, it has been shown to correctly measure changes in ozone. This was demonstrated by the high positive correlation between the AIRS and ozonesonde ozone (see Figure 4). This indicates that the data are useful for studies where the variability is important rather than the absolute value of ozone. This is a positive outcome in view of the current ozone profile being a by-product of the temperature and the channel selection has been optimized for total ozone. An improved retrieval algorithm and channel selection to optimize the ozone profile in the UTLS region will enhance the value of this data set in research applications and is in development for version 5.

Acknowledgments. We would like to thank the 3 anonymous reviewers who have helped to improve this paper. KPM would like to acknowledge the University of Canterbury and the National Institute of Water and Atmospheric research (NIWA) for PhD funding.

References
Bodeker, G. E., I. S. Boyd, and W. A. Matthews (1998), Trends and variability in vertical ozone and temperature profiles measured by ozone-


