Geophysical Trends inferred from 20 years of AIRS infrared global observations

S. DeSouza-Machado[1], L. Larrabee Strow[1,2], R. Kramer[3]

[1]JCET/GESTAR2, University of Maryland Baltimore County, Baltimore, Maryland
[2]Dept of Physics, University of Maryland Baltimore County, Baltimore, Maryland
[3]NOAA GFDL, Princeton New Jersey

Key Points:

- The 20+ year radiance record of NASA's AIRS sounder contains detailed vertical information about changes in geophysical parameters.
- A novel method retrieves geophysical trends from the radiance record, using stable channels in radiance space and straightforward *a-priori*.
- Comparisons are made to trends from monthly Numerical Weather Prediction reanalysis fields and L3 operational data products.

Corresponding author: Sergio DeSouza-Machado, `sergio@umbc.edu`

Abstract

Daily spectral radiance observations by NASA's Atmospheric Infrared Sounder contain detailed information about surface and atmospheric temperature and water vapor. We obtain climate geophysical trends from 20 years (20202/09-2022/08) of AIRS observations using a novel method operating mostly in radiance space. The observations are binned into 3 × 5 degree tiles using 16 day intervals, after which nominally clear scenes are selected for each tile to construct the spectral radiance time series. De-seasonalized spectral trends are then obtained, which are inverted using a physical retrieval to obtain geophysical trends. This approach is distinct from traditional use of radiances whereby trends are generated after operational retrievals or assimilation into Numerical Weather Prediction models. Our approach rigorously ties the derived geophysical trends to the observed radiance trends, using far fewer computational resources and time. The retrieved trends are compared to trends derived from ERA5 and MERRA2 reanalysis model fields, and NASA Level3 AIRS v7 and CLIMCAPS v2 data. Our retrieved surface temperature trends agree quite well with ERA5, CLIMCAPS and the GISS surface climatology trends. Atmospheric temperature profile trends exhibit some variability amongst all these data sets, especially in the polar stratosphere. Water vapor profile trends are nominally similar among the data sets except for the AIRS v7 which exhibits drying trends in the mid troposphere. Spectral closure between observed trends and those computed by running the reanalysis and NASA L3 monthly fields though a radiative transfer code are discussed, with the major differences arising in the water vapor sounding region.

Plain Language Summary

The current generation of infrared sounders, designed for weather forecasting purposes, have been in orbit around the Earth for a long enough time to enable anomaly and trending studies for climate purposes. Traditionally their daily obtained radiance data has been used for operational atmospheric state retrievals, or assimilation into Numerical Weather Prediction models, after which climate anomaly studies are made. In this paper we use the raw radiance spectral data to form radiance anomalies and trends, after which we do a one step atmospheric state retrieval. This novel approach has the benefit of using only stable channels together with easily understood assumptions and well tested retrieval algorithms to do the trend or anomaly geophysical retrieval, which has full error characterization.

1 Introduction

NASA's Atmospheric Infrared Sounder (AIRS) became operational in September 2002, as the first of the new generation of low noise, high stability hyperspectral sounders, making Top of Atmosphere (TOA) radiance observations at a typical 15km (at nadir) horizontal resolution. Follow on instruments with similar characteristics and abilities include the ESA's Infrared Atmospheric Sounding Interferometer (IASI) and NOAA's Cross Track Infrared Sounder (CrIS), operational since June 2007 and March 2012 respectively. The latter two already have follow on missions planned till the 2040s, and together these three sounders will provide scientists with a 40 year high quality, near continuous observational dataset for climate anomaly and trending studies.

Infrared radiances contain a wealth of information, including but not limited to surface temperature, atmospheric temperature and water amount, and mixing ratios of greenhouse gases such as carbon dioxide CO_2, CH_4 and N_2O. Measurements by visible imagers which have $\sim$ 1 km horizontal resolution or better (King et al., 2013) suggest global cloud free fractions of $\sim$ 30%, but the 15 km footprint of typical sounders means at most 5% of the hyperspectral observations can be considered "cloud-free." Current operational NASA L2 products use the method of cloud clearing on observed radiances in partly cloudy scene conditions before doing the geophysical retrieval. The cloud clearing method solves for an estimate of clear column radiances by examining adjacent Fields of View (FOVs) to estimate the cloud effects on observed allsky radiances, assuming any differences are solely due to different cloud amounts in each FOV, and significantly increases geophysical retrieval yields (to about 50-60%) (Smith & Barnet, 2023). This does introduce increased noise in the cloud cleared radiances of the lower atmosphere sounding channels; in addition the subsequent retrieval depends on the first guess (which is a neural net for AIRS v7 and MERRA2 reanalysis for CLIMCAPS v2). The reader is referred to (Susskind et al., 2003; Smith & Barnet, 2020, 2023) for more details.

In this paper we work directly in radiance space and form either anomalies or trends from the underlying well characterized and understood radiances (Strow & DeSouza-Machado, 2020), in order to do a geophysical trend or anomaly retrieval. The work presented here, once the averaged/sorted data is available, can be processed in hours to days, and can be duplicated by small research groups with ease. Moreover, our novel approach has zero temperature *a-priori* and minimal water vapor *a-priori*. This completely sidesteps time variability and the accuracy of the *a-priori* which causes errors in the retrievals, and ensures our work examines trends directly inferred from the radiances versus those from traditional methods. This leads to more unbiased results that directly highlight the conditions (for example stratospheric water vapor) where the sensor has limited sensitivity.

The approaches used in this work are therefore very different than climate anomalies or trends from reanalysis products or traditional Level 2 retrievals, neither of which are tailored for climate trends. Reanalysis products assimilate individual sensor scenes from many different instruments, and may have discontinuities as different instruments come online or go offline. Traditional Level 2 (and Level 3 products derived from Level 2) retrieve the atmospheric state for individual scenes (or effective cloud-cleared radiance derived from a 3x3 grid of individual scenes). Both reanalysis and Level 2 products require large computational resources, that preclude full dataset re-processing to help fully understand trends. A main characteristic of traditional L2 retrievals is the requirement for a good *a-priori* state for each inversion, making errors in the *a-priori* difficult to distinguish from true variability in the data, especially with regard to trends.

The stability and accuracy of the AIRS instrument is documented in recent work on analyzing 16 years of AIRS radiance anomalies over cloud-free ocean (Strow & DeSouza-Machado, 2020). Geophysical retrievals on the anomalies yielded CO_2, CH_4, N_2O and surface temperature time series that compared well against in-situ data from NOAA Global Monitoring Laboratories (GML) trace gas measurements and NOAA Goddard Institute of Space Studies (GISS) surface temperature data respectively. A significant difference between this paper and (Strow & DeSouza-Machado, 2020) is the nominally clear scenes used in this paper are selected uniformly from all over the Earth, while the clear scenes in the latter were zonal averages which were sometimes concentrated in certain regions.

In this paper we expand upon our initial zonal clear sky analysis, to derive geophysical trends from 20 years (September 2002 - August 2022) of AIRS measurements over $\sim 3 \times 5$ degree tiles covering the Earth, chosen such that the number of observations in each tile is roughly equal. An important concept introduced is spectral closure, whereby the observed clear sky spectral radiance trends are compared to spectral trends produced by running the monthly reanalysis or official NASA retrieved AIRS L3 products through an accurate clear sky radiative transfer code; close agreement in different sounding regions (such as 640-800 cm^{-1} for temperature and CO_2, 1350-1640 cm^{-1} for water vapor, 1000-1150 cm^{-1} for O_3) between the computed and actual observed spectral trends imply that trends from those geophysical parameters used in the computations are realistic while disagreement suggests otherwise. A companion paper will utilize the geophysical trend results to derive Outgoing Longwave Radiation (OLR) trends and non-local clearsky feedback parameters. Nominally clear scenes for each tile are picked out using a quantile approach; from the time series, radiances trends are made over the entire Earth, from which geophysical trends are retrieved.

Observed infrared spectral trends from AIRS has already been a focus of earlier work by (Huang et al., 2023) who studied a slightly shorter time period (2002-2020) while (Raghuraman et al., 2023) converted the radiances to Outgoing Longwave radiation (OLR), but neither study involve retrievals from spectral trends to geophysical trends. Instead they convert various model trends (such as ERA5) to spectral trends and compare against the observed spectral trends. Our earlier work shows we can accurately account for the effects of GHG forcings Strow et al. (2021). In this paper we remove these GHG forcings from the observed AIRS spectral trends to concentrate on atmospheric temperature and water vapor and surface temperature, while the papers by (Huang et al., 2023; Raghuraman et al., 2023) include the GHG forcings in the model generated spectral trends. Another noteworthy examination of the time evolution of high spectral resolution infrared radiances (converted to spectral outgoing longwave radiation (OLR) fluxes) by Whitburn et al. (2021) covered 10 years (2007-2017) of IASI observations. They confirmed that the IASI-derived fluxes agreed well with increases in GHG gas concentrations and El-Nino Southern Oscillation (ENSO) events within that time frame. A more recent paper (Roemer et al., 2023) used the 10 year IASI data to derive allsky longwave feedback spectral components (water vapor, CO_2, window, ozone) and total values, while also estimating clear sky feedback values.

We will refer to our results as the AIRS Radiance Trends (AIRS_RT). Comparisons are made against monthly output from the European Center for Medium Weather Forecast fifth generation reanalysis (ERA5) (Hersbach et al., 2020) and NASA's second generation Modern-Era Retrospective analysis for Research and Applications (MERRA2) (Gelaro & Coauthors, 2017), and also against the official monthly AIRS L3 products which are AIRS v7 L3 (Susskind et al., 2014; Tian et al., 2020) and CLIMCAPS v2 L3 (Smith & Barnet, 2019, 2020). Detailed geophysical trends and spectral closure studies are presented for the averaged ascending (daytime (D)) and descending (nightime (N)) trends; the appendix briefly discusses separate D and N trends.

2 Datasets used in this study

Three main types of datasets are used in this study. The first is the AIRS L1C radiance dataset we analyzed for this paper, which has both daytime (D) and nightime (N) (ascending and descending) views of the planet. Second is the monthly operational L3 retrieval data, which are the AIRS v7 and the CLIMCAPS v2 products, also separated into D/N data. Finally we also compared to trends from ERA5 and MERRA2 monthly reanalysis model fields. The ERA5 monthly dataset is available in 8 averaged time steps, so we match to the average AIRS overpass times and compute (D/N) data over the 20 years, while MERRA2 monthly model fields are only available as one time step; included here for completeness we mention the NASA GISS surface temperature dataset, which like MERRA2 is only available as a monthly mean. This means four of the datasets : AIRS_RT (from AIRS L1C), AIRS L3 and CLIMCAPS L3, and ERA5 are separable into D/N, while the other two (MERRA2 and GISS) are only available as a diurnal averaged value. We describe these datasets in more detail below.

2.1 The AIRS instrument and L1C dataset

The Atmospheric Infrared Sounder (AIRS) on board NASA's polar orbiting EOS/Aqua platform has 2378 channels, covering the Thermal Infrared (TIR) spectral range (roughly 649-1613 cm^{-1}) and shortwave infrared (2181-2665 cm^{-1}). The full widths at half maximum satisfy $\nu/\delta\nu \sim 1200$. The (spectral dependent) noise is typically ≤ 0.2K. The original L1b radiance dataset suffers from spectral gaps and noise contaminated data as detectors slowly fail. These limitations are addressed using a 2645 L1c channel dataset, where spectral gaps and some of the noise "pops" are filled in using principal component reconstruction (Manning et al., 2020) and is the dataset used to subset radiances analyzed in this paper. However we note that the results described in this paper used only the actual observed radiances in pristine, stable channels described in (Strow et al., 2021) and none of the synthetic channels. The Aqua platform is a polar orbiting satellite with 1.30 am descending (night time over equator) and 1.30 pm ascending (daytime over equator) tracks. Each orbit takes about 90 minutes, with the 16 passes yielding almost twice daily coverage of the entire planet. About $\sim$ 3 million AIRS spectral observations have been obtained daily since AIRS became operational in late August 2002. The instrument has provided data almost continuously since then though there have been some shutdowns (each spanning a few days) such as during solar flare events.

In this paper we use the re-calibrated 2645 channel L1C radiance data (Strow & DeSouza-Machado, 2020) instead of the 2378 L1B data. 20 years (spanning September 1, 2002-August 31, 2022) of AIRS L1C radiance data is gridded into 4608 tiles covering the Earth : 72 longitude boxes which are all 5°in width, and 64 latitude boxes which are approximately 2.5°in width at the tropics but wider at the poles to keep the number of observations per 16 day intervals (which is the repeat cycle of the AIRS orbit on the Aqua satellite) roughly the same. This way there are $\sim$ 12000 observations per 16 days per tile, which are roughly equally divided between the ascending/daytime (D) and descending/nightime (N) tracks. In this paper we discuss results for both the ascending and descending tracks using a retrieval based on the longwave (LW) and midwave (MW) regions of the spectrum (640-1620 cm^{-1} or 6-15 μm).

In this paper our trend retrievals use only the AIRS channels are stable in time, as quantified in (Strow et al., 2021). For example the shortwave (SW) channels are drifting at a higher rate than the LW/MW channels, which can lead to incorrect surface temperature rates, and are avoided in this paper. Similarly there are are many channels in the LW and MW whose detectors are drifting in time, and which are also not used here. For example there are some higher wavenumber (shorter wavelength) channels past the ozone band which have a significant drift in time, possibly due to changes in the polarization of the scan mirror coating with time. Therefore compared to other AIRS operational products used in this paper, our results use channels that are demonstrated to have high stability (Strow et al., 2021). We do note that some of the observed drifts in the AIRS channels stabilized after 6 years, so their impact is reduced when looking at 20 year trends.

2.2 Reanalysis Model fields

The ERA5 fifth generation reanalysis product from the European Center for Medium Range Weather Forecasts is freely available on monthly timescales from the Copernicus Climate Data Store. This monthly dataset is output at 37 pressure levels at 0.25°horizontal resolution (Hersbach et al., 2020), which is further subdivided into eight 3-hour averages per month (corresponding to 00,03,06,...21 UTC). For each month from September 2002-August 2022 we downloaded the surface temperature and pressure fields, as well as atmospheric temperature, water vapor and ozone fields. These are then colocated to each tile center using 2D spatial interpolation, as well as time interpolated according to the average AIRS overpass time as a function of month. From the resulting monthly timeseries of reanalysis model fields for each tile, we generated (a) thermodynamic trends for surface temperature, air temperature, water vapor and ozone model fields (b) a 20 year average thermodynamic profile in order to produce jacobians for the linear trend retrievals (c) by using the model fields as input to the clear sky SARTA radiative transfer code (Strow, Hannon, DeSouza-Machado, et al., 2003) a monthly time series of clear sky radiances for each tile was generated, from which we could compute radiance trends. We did this for both the ascending and descending datasets.

The MERRA version 2 (MERRA2) re-analysis used in this paper is the second generation (Gelaro & Coauthors, 2017) product from NASA's Global Modeling and Assimilation Office. The monthly data we use is available on 42 pressure levels at a horizontal resolution of 0.5°× 0.625°, but only one monthly mean diurnally averaged output is available per month. Similar to the ERA5 output, we colocated the MERRA2 surface temperature, atmospheric temperature, water vapor and ozone fields to our tile centers for each month starting September 2002 in order to produce a time series of radiance and model output, from which radiance and thermodynamic trends could be computed for comparisons against other datasets in this study; similar to above we also generated a monthly time series of clear sky radiances for each tile, from which we could compute clear sky radiance trends based on MERRA2.

The NASA Goddard Institute of Space Studies (GISS) surface temperature data v4 surface temperature data (2023, 2005; Lenssen et al., 2019) is a monthly dataset based primarily on near surface temperatures land stations, and data from ships and buoys. As with MERRA2 we obtained one monthly mean dataset per month, which we could not separate into descending (N) or ascending (D) tracks.

2.3 AIRS L3 Products

NASA routinely produces two retrievals from the AIRS L1C data observed each day, which are AIRS v7 (Susskind et al., 2014; Tian et al., 2020) and CLIMCAPS v2 (Smith & Barnet, 2019, 2020). Both use the cloud clearing process but there are significant algorithmic differences; in particular the AIRS v7 product is initialized by a neural net, while CLIMCAPS uses MERRA2 for its initialization. The L2 products are then individually turned into L3 monthly products, for both the ascending (daytime) and descending (nighttime) data. The timeseries of thermodynamic profiles were used as input to the clear sky SARTA RTA to generate radiances, after which radiance trends and thermodynamic trends are also produced.

2.4 Other L3 Products

The Microwave Limb Sounder (MLS) monthly binned water vapor (H2O) mixing ratio dataset (Lambert et al., 2021), which contains data at spatial coverage ±82° latitude, at a spatial resolution of 4°× 5°and useful vertical range between 316 and 0.00215 hPa was used in this paper to improve retrieval trends in the upper atmosphere.

3 Filtering the Observational Data for clear scenes

Here we discuss the "clear scene" selection from all the observed data stored for each of the 72 × 64 tiles. Ideally we would prefer to use a MODIS cloud fraction product (1 km) colocated to the 15 km AIRS footprints, but this is presently unavailable. Our earlier work used an uniform clear flag over ocean (Strow et al., 2021) which will not work well over land because of surface inhomogeneity. In this section we discuss an alternative clear filter based on the hottest 10 percent of AIRS observations that are present inside any 16 day tile, over any location.

3.1 Observed BT1231 Distributions

The radiances measured in thermal infrared window region (800-1000 cm^{-1} and 1100-1250 cm^{-1}) are dominated by the effects of the surface temperature, water vapor continuum absorption and cloud/aerosol effects. The effects of water vapor continuum absorption is largest in hot and humid tropical scenes (depressing the observations relative to surface temperatures by about 5-6 K, which reduces to about 2 K at ± 50°) and is almost negligible for cold, dry scenes (less than 1 K). Scattering and absorption by liquid and ice clouds also affects the window region (Deep Convective Clouds can depress the window channel observations by as much as 100 K relative to surface temperatures). For each tile, we use the 1231.3 cm^{-1} observation as our representative window channel (AIRS L1C channel ID = 1520), as it is minimally impacted by weak water vapor lines. Changed to Brightness temperature (BT) the observation in this 1231.3 cm^{-1} channel (BT1231) therefore serves as a measure for the cloudiness of an observation : if there are no or low or optically thin clouds, it will effectively measure the surface temperature, but as the clouds get thicker and higher, it will measure the cold cloud top temperatures. For any tile during any 16 day observation periods, we can compute quantiles $\mathcal{Q}$ based on the observed BT1231 to screen between cloudy and partially clear scenes. We chose different BT1231 quantiles (so quantile $\mathcal{Q}$0.XY will have a numerical value $BT1231_{\mathcal{Q}0.XY}$ associated with it) and show below the data contained between $\mathcal{Q}$0.90 and $\mathcal{Q}$1.00 can be considered "almost free of clouds."

Figure 1 shows all the BT1231 observations for a chosen 16 day timestep in the form of a zonally averaged histogram (normalized probability distribution functions (PDFs)), with latitude on the vertical axis and BT1231 on the horizontal axis. The colorbar is the PDF value, and we used data spanning August 27, 2012 - September 11, 2012 which is approximately half way through the 20 year AIRS mission dataset used in this paper. The curves show the zonally averaged BT1231 values of the minimum ($\mathcal{Q}$0.00) in dark cyan, mean (thick red), median ($\mathcal{Q}$0.50 in orange), maximum ($\mathcal{Q}$1.00 in light cyan); also shown are a handful of other zonally averaged BT1231 values, for example $\mathcal{Q}$0.80, $\mathcal{Q}$0.90 (thick black curve), $\mathcal{Q}$0.95 and $\mathcal{Q}$0.97. The distributions are skewed to the left (negative skewness), as confirmed by the mean being less than the median. We also point out that even $\mathcal{Q}$0.80 sees much of the surface from the southern tropics to the northern polar region. The 220 K horoizontal axis cutoff means we do not see the very cold (190 K) observations over the winter Antarctic.

The figure shows the expected qualitative features, for example (1) the tropical PDFs peak at around 295 K, but show some warmer observations, as well much colder observations (below 230 K) corresponding to Deep Convective Clouds (DCC); this gives a dynamic range of almost 100 K at the tropics (2) the BT1231 observed over the Southern Polar (polar winter) regions are much colder than the BT1231 observed over the Northern Polar (polar summer) regions and (3) the reddish peaks in the 30°N - 40°N are a combination of the marine boundary layer (MBL) clouds and warmer summer land temperatures. Figure 1 shows on average the cloud effect at the tropics is an additional modest 20 K (difference between $\mathcal{Q}$0.90 and $\mathcal{Q}$0.50) compared to the 100 K dynamic range. This is because the cloud fractions and cloud decks in the individual observations have effectively more clouds (with larger cloud fraction in the FOV) lower in the atmosphere than higher up; the net effect is that in the window region the atmosphere is on average radiating from the lower (warmer) altitudes, and so spectra whose BT1231 values are larger than $BT1231_{\mathcal{Q}0.80}$, see much of the surface emission as well.

We now use the above plots to select "almost clear" scenes. For any one tile, we define set $\Psi_{0.XY}$ to have all observations i whose BT1231 lies between quantiles $\mathcal{Q}$0.XY and $\mathcal{Q}$1.00, $\{i \mid BT1231_{\mathcal{Q}0.XY} \leq BT1231(i) \leq BT1231_{\mathcal{Q}1.00}\}$. In what follows Q0.XY is the radiances averaged over all the observations i which are in the set $\Psi_{0.XY}$, namely

$$r_{Q0.XY}(\nu) = \frac{1}{N_{0.XY}} \sum_{i \in \Psi_{0.XY}} r_i(\nu) \tag{1}$$

where $r_i(\nu)$ are the $N_{0.XY}$ individual observations in set $\Psi_{0.XY}$. In this section we only use the $\nu = 1231$ cm^{-1} channel, but in later sections we easily form averages for all 2645 channels, at any 16 day time step for any tile.

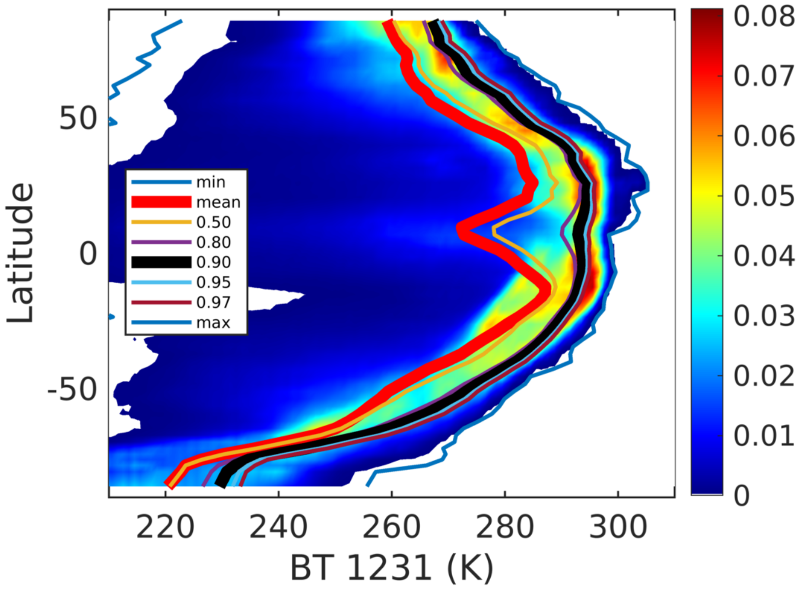

Figure 1. Zonally averaged BT1231 normalized histograms (probability distribution functions) as a function of latitude and temperature bin, for an 2012/08/27 - 2012/09/11 timespan (colorbar) and quantiles (curves). The thick black curve is the $\mathcal{Q}0.90$ quantile (and above) used in this paper, and is very close to the maximum $\mathcal{Q}1.00$ quantile.

We tested different quantile sets $\Psi_{0.XY}$ to see which one can reliably be considered to provide a nominally "cloud free" global dataset, and chose the Q0.90 average (ie defined as averaged over the $\Psi_{0.90}$ set) as the one to use for the rest of this paper, unless explicitly stated otherwise. The tests primarily involved comparisons to scenes produced by the uniform/clear sky filter described in (Strow & DeSouza-Machado, 2020) for the same August 27, 2012 - September 11, 2012 sixteen day timespan. This latter filter selects clear scenes by both testing for uniformity (to within 0.5 K) across a 3 × 3 grouping of AIRS scenes and also using a criteria that the observed window channel observations should be within ± 4 K of clear-sky simulations using thermodynamic parameters supplied by reanalysis models. The results are shown in the left hand plot of Figure 2, plotted on a 1°× 1°grid. We note in this plot the uniform/clear scenes that are plotted are limited to those over ocean, and for solar zenith less than 90 °(daytime), which automatically filtered out many of the views over the (wintertime) Southern Polar region. Immediately apparent are the gaps produced by the uniform/clear filter *e.g.* in the Tropical West Pacific or off the western coasts of continents where there are clouds. The gaps can be changed by *e.g.* changing the 4K threshold to allow more or fewer scenes through the filter.

The center plot shows for all tiles, the daytime scenes selected by the Q0.90 filter for the same time period, on the same 1°× 1°grid. Compared to the left hand plot, the spatial coverage is almost complete, as the Q0.90 average always has the hottest 10% of the observations. At this 1°resolution, used for comparison with the uniform/clear grid filter described in the previous paragraph, gaps are seen in regions where for example the local topography means observations over mountains would be colder than the surrounding coastal or plain regions. This is not a concern since zooming back out to the coarser 3°× 5°tile resolution, will include Q0.90 data for the quantile and trending analysis.

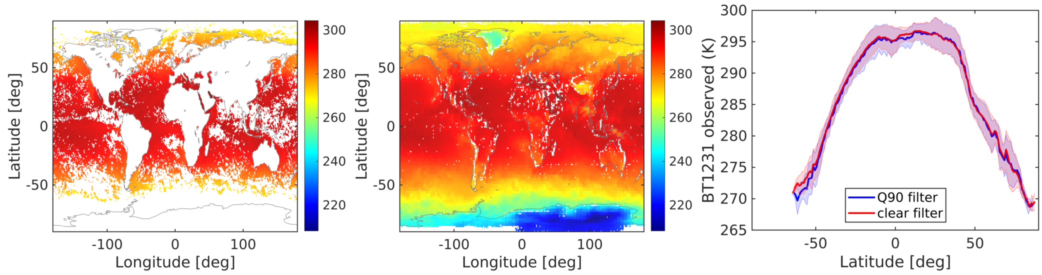

Figure 2. Clear scenes for the same 2012/08/27 - 2012/09/11 timespan selected by (left) an uniform/clear sky filter and (center) the Q0.90 average described in this paper. The right hand plot shows the mean (over ocean) observed BT1231 as a function of latitude, for the two selections; the difference is about 0 K $\pm$ 1 K in most region except in the southern midlatitudes where the Q0.90 average produced scenes that were about 1 K cooler on average.

To compare the mean observations we remove the over-land and over-polar region data from the center plot. The right hand plot shows the mean observed BT1231 from the 1°× 1°grid from the uniform/clear sky filter as a function of latitude, compared to the 1°× 1°grid from the Q0.90 scenes. The difference between the uniform/clear versus Q0.90 average is within about 0.25 K $\pm$ 1 K across the southern tropics to the northern midlatitudes, though the bias rises to about 1 K by about -50°S. We consider this an acceptable difference, as we could tune the thresholds for the uniform/clear filter to *e.g.* change the areal coverage and/or number of clear scenes and hence comparisons to the Q0.90 scenes.

The results presented in this section have been checked for robustness, using other 16 day intervals spanning the four seasons. We conclude that for any 16 day timestep the radiances used in the Q0.90 average (a) produces almost complete spatial coverage of the Earth, (b) selects scenes whose average BT1231 is very close to the average BT1231 from scenes selected using an uniform/clear filter (c) trends from that quantile typically differ by less than $\pm$ 0.002 K yr^{-1} from the other quantiles and (d) this selection produces spectral trends which compare well against those obtained from the quality assured binned AIRS CCR data record (Manning, 2022). Together these imply the Q0.90 average is an acceptable proxy for "clear scenes". For the remainder of the paper we therefore consider Q0.90 as consisting of nominally clear observations whose BT1231 lies between the 90th quantile and hottest observation. Our retrievals using this $Q0.90 \rightarrow Q1.00$ averaged dataset (shortened to Q0.90) is referred to as AIRS_RT in what follows.

3.2 Observed trends from the Q0.90 Quantiles

Having selected the Q0.90 observations, for each tile the average radiance per 16 day interval is computed. With two sixteen day periods not available (Aqua platform or AIRS shutdowns during *e.g.* solar flare events) this gives a total of 457 time steps over 20 years. Anomalies are formed from this time series, and then de-seasonalized to give the spectral radiance trends and error estimates (Strow & DeSouza-Machado, 2020) using Matlab *robustfit*:

$$r^{\text{16 days}}_{\text{observations}}(t) \;\sim r_{\text{fit}}(t) = r_o + a_1 t + \sum_{i=1}^{4} c_i sin(n2\pi t + \phi_i) \tag{2}$$

with a_1 and its associated uncertainty, both converted to brightness temperature (BT), being the trends in K yr^{-1}. Using sub-harmonics in the fit did not produce any noticeable change in the AIRS_RT retrievals (described below).

The left panel of Figure 3 shows the descending orbit (nighttime) 20 year (September 2002- August 2022) global averaged spectral observations for the five quantiles mentioned above. We note the spectra in most of the plots in this section are weighted by the *cosine(latitude)* of the tiles, unless otherwise stated. In addition we only show the 640-1640 cm^{-1} region, and ignore the shortwave 2050-2750 cm^{-1} region since the AIRS SW channels are drifting relative to the LW (Strow & DeSouza-Machado, 2020). Spectral averages constructed from Figure 1 would have this same behavior, namely that in the window region the mean spectrum of data populating the warmer quantiles (Q0.80, Q0.90, Q0.95, Q0.97) as defined in Equation 1 are on the order of a Kelvin apart, and have about half/quarter that difference in the optically thicker regions dominated by H_2O and/or CO_2 absorption respectively.

The right hand panel of Figure 3 shows (top) the trends and (bottom) the 2σ trend uncertainties for these quantiles, in K yr^{-1}. We emphasize that the top right panel shows that the spectral trends for the quantiles lie almost on top of each other; the difference between the Q0.50 and other trends is at most about $+0.003$ K yr^{-1} (out of a 0.02 K yr^{-1} signal) in the window region (and about $+0.0045$ K yr^{-1} in the troposphere temperature sounding channels), or less than 10%. Similarly the largest trend uncertainty in the bottom panel is for Q0.50. This implies that clouds effects in the infrared do produce the largest variability (blue curve) but on average for the infrared are not changing much, so the $+0.022$ K yr^{-1} window region trends are dominated by surface temperatures changes and to a lesser extent by water vapor changes.

TOA radiances in the 15 um (700-800 cm^{-1}) region are impacted by two effects (a) the increased optical depths due to increasing atmospheric CO_2 leads to atmospheric emission from higher altitudes/lower temperatures, resulting in almost a -0.06 K/year signal for the troposphere, and (b) the atmospheric temperature increases (again about $+0.02$ K yr^{-1}). Also of interest is the trends in the stratosphere (650-700 cm^{-1}) changes which consists of a stratospheric cooling signal (negative) and emission higher up due to increased CO_2; combining to give a net zero effect over 20 years, also seen in (Raghuraman et al., 2023). The H_2O signal is evident in the 1400-1625 cm^{-1} region, and is only slightly positive; in other words, increasing temperatures have led to increased atmospheric amounts of H_2O, and the water vapor feedback has reduced the amount of outgoing flux in that region. By extension, this can only be expected to have happened in Far Infrared (10-650 cm^{-1}) spectral regions affected by water vapor, and cannot be wholly confirmed as current sounders do not make direct measurements in that region. In the near future it is anticipated the Far Infrared Outgoing Radiation Understanding and Monitoring (FORUM) mission (Palchetti et al., 2020) will provide data to fill in this important observation gap.

4 Spectral closure : comparisons between observed and simulated spectral trends

Previous work (Strow & DeSouza-Machado, 2020) has demonstrated that the radiances from AIRS are climate quality, if one restricts the channel set to the ~ 450 channel set that is largely immune to nonphysical drifts (Strow et al., 2021). In this section we describe a way to test the quality of the monthly thermodynamic output from reanalysis and/or L3 products which are all in geophysical space, against the AIRS L1C observational data which is in radiance space. This is accomplished by geolocating the monthly (ERA5) surface temperature, air temperature, water vapor and ozone fields to tile cen-

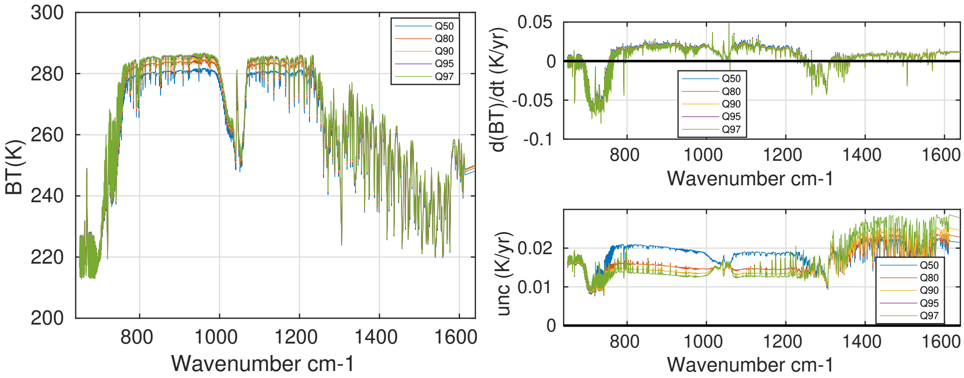

Figure 3. 20 year trends from different observation quantiles. The left hand panel shows the mean globally averaged BT observations from 20 years of AIRS data, for quantiles Q0.50,0.80,0.90,0.95,0.97 as described in the text.The right hand panel shows (top) the globally averaged trends for those different quantiles and (bottom) the spectral uncertainty in the trends. The nighttime (descending) trends are shown in these plots.

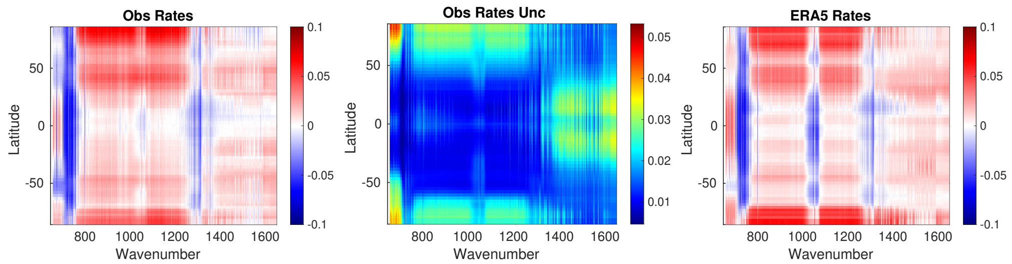

Figure 4. 20 year zonally averaged spectral brightness temperature trends (in K yr^{-1}) for (left) AIRS Q0.90 observations and (right) clear sky simulations using ERA5 monthly model fields. The center panel shows the AIRS Q0.90 spectral uncertainties. The ERA5 simulations included linear trends of CO_2, CH_4 and N_2O, while the O_3 trends in ERA5 are from the reanalysis itself.

ters as described in Section 2.2, which are then input and run through the SARTA fast model (Strow, Hannon, DeSouza-Machado, et al., 2003), for the entire 20 years. Spectral radiance trends were then computed from these time series of (clear sky) spectral radiances. The conversion of L3 retrieval and NWP reanalysis trends to a radiance time series, provides a rigorous check of their accuracy against the observed AIRS L1C radiance trends which are validated to be highly accurate.

The simulations included realistic column linearly-increasing-with time mixing ratios for CO_2, CH_4 and N_2O for the ERA5 spectra, as well as land or ocean surface emissivity co-located to tile centers together with view angles of about 22°, which is the average view angle of the tiled observations. From these the ERA5 spectral trends were derived similarly to what was described above for the AIRS observation spectral trends.

Figure 4 shows the AIRS observed Q0.90 (nominally clear) descending (night) zonally averaged results in K yr^{-1} in the left panel, and the zonally averaged simulated clearsky (without clouds) spectral trends (also in K yr^{-1}) from monthly ERA5 fields in the right panel. The center panel shows the spectral trend uncertainties from the observations, also in K yr^{-1}. In the next section we derive geophysical trends from these (AIRS observed) spectral trends, and the similarities/ differences in geophysical trends between observations and models/operational data can be partially understood from the similarities/differences in the spectral trends. For example, the H_2O sounding region (1350-1600 cm^{-1}) of the left and right panels of Figure 4 shows roughly similar (positive) spectral trends in the tropics and mid-latitudes; there are some slight differences in the high altitude channels (1450-1550 cm^{-1} region). The following sections shows that this results in subtle differences in the tropospheric water vapor trends. Observations and simulations both have positive dBT/dt in the 800-960,1150-1250 cm^{-1} region, indicating surface warming; however the ERA5 simulation show more warming in the southern polar regions than do the AIRS observations. Note the mean warming in the tropics for both observations and ERA5 simulations is less than that in the mid-latitudes, and the polar regions show the largest overall change in brightness temperature in the window region. Large differences are seen in the 10 um (1000 cm^{-1}) O_3 sounding region, which are not surprising since ozone assimilation is not a primary goal of ECMWF assimilation; here we do not address these as we focus on the changes to the moist thermodynamic state. The window region trends computed using the ERA5 model are more positive in the Southern Polar region. Conversely the 640-700 cm^{-1} spectral region is positive, especially in the tropics; however the observations show a net cooling trend away from the tropics, compared to the ERA simulations. This demonstrates the importance of the model $\rightarrow$ spectral trend comparisons, given the accuracy of the AIRS observations.

The paper by Raghuraman et al. (2023) shows similar figures, but in terms of spectral OLR trends encompassing the 0-2000 cm^{-1} range, while Huang et al. (2023) shows similar plots for a slightly smaller time period (2002-2020) using the nadir L1B radiance dataset which has no or minimal frequency corrections compared to the L1C set we use in this paper. Huang et al. (2023); Raghuraman et al. (2023) and our work all show, either in radiance or OLR space, (a) the increased observed radiance in the window channels, due to surface temperature increases (b) the $\simeq$ -0.06 K yr^{-1} decrease in BT in the 700-750 cm^{-1} troposphere sounding region, which is due to the CO_2 amounts increasing; we also see differences in the signs of the BT changes in the 650-700 cm^{-1} stratospheric CO_2 and temperature channels for some latitudes between AIRS_RT observations and ERA5 simulations (c) increases in the 1350-1640 cm^{-1} water vapor sounding region seen in Figures 3 and 5, and (d) the 1280-1340 cm^{-1} decreases are due to CH_4 increases.

4.1 Sample spectral closure comparisons using other monthly products

Here we follow the earlier work of (Huang et al., 2023) and convert the ERA5 monthly model fields to spectral radiances, after which we compute spectral trends for comparison to AIRS observations. Spectral closure calculations for the entire 20 year timeseries were also generated for the monthly MERRA2 model fields, as well as the monthly AIRS v7 L3 and CLIMCAPS L3 retrieved data products. Again only the monthly thermodynamics and surface temperature fields for all 72 $\times$ 64 tiles were used in these SARTA runs, with GHG changes added in for each timestep as described above. Spectral trends were then computed using Equation 2.

We chose just one limited example here to illustrate the power of this approach for diagnosing which dataset is more accurate, given that the AIRS spectral trend accuracy is already established. Water vapor is highly variable in space and time, meaning water vapor retrievals using hyperspectral sounders radiances differ most from NWP forecasts, in particular because of the typical $\pm$ 90 minute difference between observation and forecast, and is where these sounders typically provide the most information. Figure 5 show the globally averaged brightness temperature trends (in K yr^{-1}) in the 1350 - 1650 cm^{-1} water vapor sounding region. The blue curve shows the trends from the AIRS

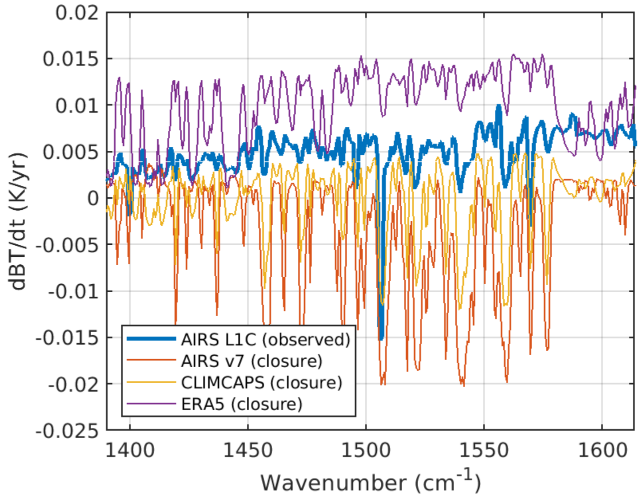

Figure 5. Globally averaged spectral trends in the water vapor sounding region : AIRS L1C observations (blue) compared to spectral closure from the standard monthly AIRS L3 retrievals (red) and CLIMCAPS L3 (yellow) and from monthly ERA5 simulations (yellow). The reconstructed AIRS_RT trends very closely match the AIRS L1C observations and are not shown here.

observations used in this paper, while spectral trends constructed from the AIRS L3/ CLIMCAPS L3 retrievals are in red/yellow and the ERA5 model fields are in purple. The AIRS observations and ERA5 constructed spectral trends are positive in this region, while the AIRS L3 and CLIMCAPS L3 trends are obviously different, being negative in this water vapor sounding region. The subtle differences in these spectral trends arise from differences in the geophysical trends between observations and the models themselves, and will be addressed in the following sections, where the retrieved and model surface temperature, and atmospheric temperature and water vapor geophysical trends will be compared and discussed.

5 Testing the variability of representative points from NWP reanalysis

Each sixteen day 3°× 5°tile contains ∼ 12000 observations, which means for each tile about 600 daytime and 600 nightime observations are averaged to produce the Q0.90 dataset per timestep. Conversely there are typically only ∼ 240 monthly ERA5 0.25°points per 3°× 5°tile; for 1°resolution AIRS L3 and CLIMCAPS L3 there are even fewer (15) points per tile. This low number of points means we chose a simple solution of using the grid cell closest to the center of each 3°× 5°tile for building the NWP and L3 geophysical time series. This choice is validated below using the following test to see for example how surface temperature trends would be impacted as we changed the representative point for the ERA5 model fields.

For the descending overpass we built complete sets of approximately 240 ERA5 points per tile per month; at 0.25°resolution one of these is almost certainly at the tile center. From these monthly sets, we could either directly read the tile center temperature (our default), or compute the average surface temperature per tile, or compute the average of the hottest 10% surface temperatures per tile. This was done for all 20 years (240 monthly timesteps) after which the three timeseries were trended. Over ocean the differences between all three datasets was typically -0.001 ± 0.005 K yr^{-1}, while over land the differences were about 0.001 ± 0.01 K yr^{-1}. This is to be compared to mean trends of about 0.014 ± 0.02 K yr^{-1} over ocean and 0.025 ± 0.04 K yr^{-1} over land : the spread of the ocean and land ERA5 surface temperature trends for the three methods, is much smaller than the mean trends. Given that there were far fewer re-analysis points in a grid box than tiled Q0.90 observations, coupled with the fact that choosing the 10% warmest profiles would provide an even smaller sample, we chose to use the tile center to be the representative point to co-locate the model fields.

6 Geophysical Trend Retrieval outline

6.1 Setting up the Retrieval Problem

The observed clear sky spectral brightness temperature for a tile at any time t can be modeled as

$$BT(\nu, t) = f(X(t), \epsilon(\nu, t), \theta(t)) + \mathrm{NeDT}(\nu) \tag{3}$$

where the state vector $X(t)$ has the following five geophysical state parameters : (1) surface temperature (ST), (2) atmospheric temperature profile T(z), (3) water vapor profile WV(z), (4) ozone profile O3(z) (5) greenhouse gas forcings (GHG) due to CO_2, CH_4 and N_2O changing as a function of time t and $f(X(t), \epsilon, \theta, \nu)$ is the clear sky radiative transfer equation for channel center frequency ν. The spectral noise NeDT(ν) for a typical tropical "clear scene" is about 0.1 K in window region, increasing to about 1 K in the 15 μm temperature sounding channels and about 0.2 K in the 6.7 μm water vapor sounding region, but the noise will vary as a function of scene temperature. We parametrize the GHGs using single numbers (such as ppm(t) for the CO_2 column), and include the AIRS orbit and viewing angle geometry θ and the surface emissivity $\epsilon(\nu)$, while we omit forward model and spectroscopy errors. We ignore cloud scattering as well as the spatial variation of the state parameters, emissivity and scan angle geometry within a tile. Linearizing the above equation about the time averaged profile, the relationship between the observed spectral trends and desired thermodynamic trends is given by

$$\frac{d\overline{BT(\nu)}}{dt} = \frac{\partial f}{\partial \overline{X}} \frac{d}{dt}\overline{X(t)} = K(\nu)\frac{d}{dt}\overline{X(t)} + \cancelto{0}{K_{\mathrm{emissivity}}(\nu)\frac{d}{dt}\overline{\epsilon(t)}} \rightarrow K(\nu)\frac{d}{dt}\overline{X(t)} \tag{4}$$

where the matrix $K(\nu)$ is the thermodynamic jacobian (surface temperature, air temperature and trace gases) and we ignore any orbit drifts (changes to θ), instrument changes (changes to $NeDT(\nu)$) and surface emissivity ($\epsilon(\nu)$); the last assumption is investigated in a later section. The overbars on parameters X denotes this is a time average (linear trend) that we are working with, and we have converted from radiances in Equation 2 to brightness temperatures in Equations 3 and 4.

6.2 Jacobian calculations

For a typical clear sky tropical sky atmosphere, the 800 - 1200 cm^{-1} window region has surface temperature (SKT) jacobians which are about +0.5 to +0.75 K per degree SKT change and -0.75 to -0.25 K per 10% change in column water vapor. The spectral variability in these window region jacobians is primarily due to reducing water continuum absorption as you move from the 800 cm^{-1} end to the 1200 cm^{-1}; consequently the surface temperature jacobians becomes closer to unity and the column water jacobians become closer to zero as water vapor amount decreases (drier atmospheres in the mid-latitudes and polar regions). The hyperspectral channels used in this work assist in partitioning these two competing changes (though not perfectly), which we validate against other datasets in this study. As seen in Figure 4 typical magnitudes of the spectral trends on the left hand side of Equation 4 are less than about 0.1 K per year. Equation 4 is in the usual inversion form $\delta y = K\delta x$, and the Optimal Estimation Rodgers (2000) solution used to solve the anomaly time series in (Strow et al., 2021) is also used here. The noise term used for the trend retrieval $NeDT(\nu)$ is not the instrument noise since each 16 day point in our time series is averaged over hundreds of observations as earlier described; instead the uncertainty is that due to inter-annual variability in the linear trends obtained from the trend fitting in Equation 2. Examples of typical noise values are shown in the bottom right hand panel of Figure 3.

ERA5 monthly model fields at tile centers, together with time varying concentrations of GHG such as CO_2, were averaged over 20 years so jacobians could be computed. The GHG concentrations were a latitude dependent increase of about 2.2 ppm yr^{-1} for CO_2 derived from the CarbonTracker (Peters et al., 2007) (CarbonTracker CT-NRT.v2023-4, http://carbontracker.noaa.gov) model data at 500 mb. Our pseudo-monochromatic line by line code kCARTA (De Souza-Machado et al., 2018, 2020) was used with these averaged profiles to produce accurate analytic jacobians. The HITRAN 2020 line parameter database (Gordon & Rothman, 2022), together with MT-CKD 3.2 and CO_2,CH_4 line mixing from the LBLRTM suite of models (Clough et al., 2005) were used in the kCARTA optical depth database (De Souza-Machado et al., 2018). A 12 month geographical land-varying spectral emissivity database spanning one year from (Zhou et al., 2011) was used, while ocean emissivity came from (Masuda et al., 1988). The atmospheric temperature, water vapor and ozone profile jacobians, and the surface temperature and column jacobians for the GHG gases such as CO_2 and CH_4 and N_2O, were then convolved using the best estimate AIRS Spectral Response Functions (Strow, Hannon, Weiler, et al., 2003).

Tests done for this paper, together with the results in (Strow et al., 2021), established that jacobians derived from MERRA2 versus ERA5 produced no significant differences in the context of retrieved trends or anomalies done for this paper, as the uncertainty in linear trends due to inter-annual variability dominates over any uncertainty (or differences between) model fields.

6.3 Optimal Estimation Retrieval : State vector, covariance matrices and *a-priori*

Using monthly ERA5 model fields averaged over 20 years, for each of the 64 × 72 tiles we computed analytic jacobians for the following (vector) atmospheric thermodynamic variables [fractional water vapor, fractional ozone and temperature] together with (scalar) surface temperature. We retrieved fractional gas concentration trends $dfracX/dt = 1/X_{avg}(z)dX_{avg}(z)/dt$ to keep all values in the state vector at about the same magnitude. A single iteration Optimal Estimation retrieval (Rodgers, 2000) is used to simultaneously solve for the geophysical parameter trends. As in Strow & DeSouza-Machado (2020) the geophysical covariance uncertainty matrices are a combination of Tikonov and covariance regularization. The uncertainties for the covariance matrices were typically [0.1,0.25,0.45] K yr^{-1} for the surface/tropospheric/stratospheric temperature trends, and [0.04/0.02] yr^{-1} for the fractional tropospheric/stratospheric water vapor trends. Tikonov L1 regularization Rodgers (2000) also included, with the scalar factor multiplying this regularization corresponding to about 1/10 the covariance uncertainties. The spectral uncertainties used in the retrievals come from the above mentioned trend uncertainties. For completeness we note that a sequential retrieval (see for example (Smith & Barnet, 2020)) produces very similar geophysical trends.

Here we emphasize four points about our geophysical trend retrievals, which sets us apart from trends derived from other datasets. Firstly the *a-priori* trend state vector is zero (dST/dt = dT(z)/dt = dQ(z)/dt = 0) for all geophysical parameters, except for water vapor where we enforced constant (or slightly increasing) relative humidity as described below. This ensures traceability of our retrieval is straightforward especially wherever the AIRS instrument has sensitivity. For example the 300 - 800 mb water vapor trend retrievals will be based on the data only, thereby insulating us from any possible *a-priori* information from *e.g.* climatology or NWP models, unlike the operational AIRS V7 or CLIMCAPS retrievals which use first guesses based on neural net and MERRA2 respectively.

Secondly as seen in Figures 4 and 5, in the 15 μm region there is a large spectral overlap signal (-0.06 K yr^{-1}) from the increasing CO_2, which is much larger than the expected atmospheric temperature trend (0.01 K yr^{-1}). These correlations makes it difficult to jointly retrieve both temperatures changes and changes in well mixed GHGs such as CO_2. We chose to focus on retrieving temperature changes only, by spectrally removing the effects of changing CO_2, CH_4 and N_2O GHG concentrations. This was done by using the GHG trends estimated from NOAA ESRL CarbonTracker data multiplied by the appropriate GHG gas column jacobian (CO_2,N_2O and CH_4 and CFC11,CFC12) computed as described above using the averaged over 20 years ERA5 monthly profile for each tile.

Thirdly instead of using all 100 layers described in the AIRS forward model (Strow, Hannon, DeSouza-Machado, et al., 2003), we combine pairs of layers for a 50 atmospheric layer retrieval, as the AIRS radiances contain far fewer than 100 pieces of information (see *e.g.* (Maddy & Barnet, 2008; De Souza-Machado et al., 2018)).

Fourthly, modern hyperspectral infrared sounders have highest sensitivity to temperature and water vapor in the mid-tropopause; see for example the averaging kernels in (Irion et al., 2018). Using a zero fractional WV trends *a-priori* at all levels, it was fairly straightforward to obtain fractional WV(z) trends close to those from the NWP model datasets in the 300-850 mb region. In order to improve our results in the lowest layers, we enforced a constant relative humidity approximation, which is a well-known, expected behavior under global climate change (Soden & Held, 2006; Sherwood et al., 2010). This was done by ignoring the contribution due to water vapor changes in the observed BT1231 trend, and using it as an approximation for air temperature trend over ocean; this allows us to compute an estimate of how the water vapor would need to change

$$RH(T) = \frac{e}{e_{sat}(T)} \implies \delta(RH) = \frac{1}{e_{sat}(T)}\delta e - \frac{e}{e^2_{sat}(T)}\delta e_{sat}(T) = \frac{1}{e_{sat}(T)}\delta e - \frac{e}{e_{sat}(T)}\frac{L_v}{R_v}\frac{1}{T^2}\delta T \tag{5}$$

where $e, e_{sat}(T)$ are the vapor pressures and we used $e_{sat}(T) = e_{s0}e^{\frac{L_v}{R_v}\left(\frac{1}{T_o}-\frac{1}{T}\right)}$ (where L_v, R_v are latent heat of vaporization and gas constant respectively) to go from the expression in the center to the expression on the right. If we expect the change in RH to be zero then $\frac{\delta e}{e} = \frac{L_v}{R_v}\frac{\delta T}{T^2}$, where we can use $\delta T/\delta t \sim d/dtBT1231.$ to approximate the *a-priori* fractional vapor pressure rates (or *a-priori* fractional water vapor rates) between surface and 850 mb, smoothly tailing to 0 in the upper atmosphere. Subsection 7.2 has a similar discussion on a proposed method to alleviate the lack of sensitivity to upper atmosphere water vapor. Our default results in this paper are from using the MLS *a-priori*, unless otherwise stated.

6.4 Testing on synthetic trend spectra made from ERA5 Reanalysis monthly fields

We tested the retrieval code by using it on the simulated nighttime only ERA5 spectral trends, and compared to geophysical trends computed directly from the ERA5 reanalysis model. Spot checks of the spatial correlations of ERA5 fractional water vapor and temperature trends versus the trends retrieved from synthetic spectra/our retrieval algorithm, peaked at 500 mb with correlations of about 0.9, compared to 800 mb correlations of 0.80 and 0.55 for temperature and fractional water vapor trends respectively and 200 mb correlations of 0.89 and 0.69 for dT/dt, dWVfrac/dt. This is to be expected since a computation of the water vapor averaging kernels for infrared instruments for arbitrary atmospheric profiles typically shows they peak in the 300 mb - 850 mb range and

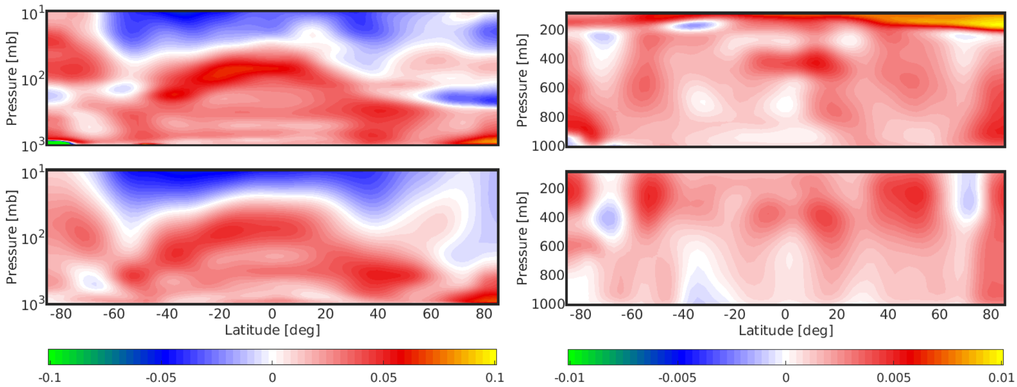

Figure 6. Comparing geophysical trends derived directly from ERA5 monthly nighttime fields (top) vs from the AIRS_RT retrieval applied to the ERA5 reconstructed spectral trends(bottom). Left panel is dT/dt (in K yr^{-1}) while rightmost panel is d(fracWV)/dt (colorbar in yr^{-1}).

decrease rapidly away from those regions; conversely the temperature averaging kernels stay relatively uniform through the free troposphere and above, though they also decrease close to the surface; see for example (Irion et al., 2018; Smith & Barnet, 2020; Wu et al., 2023).

Figure 6 shows a sample set of results using nightime ERA5 model output converted to spectral trends as described above. The top panels (A) are always the atmospheric trends computed directly from the monthly ERA5 model fields, while the bottom panels (B) are the atmospheric trends retrieved from the converted ERA5 spectral brightness temperature trends. The left most panel is the atmospheric temperature trend comparison (both in K yr^{-1}) while the rightmost panel is the fractional atmospheric water vapor trend comparison (in yr^{-1}).

It is evident from the figure that the tropospheric trends in the tropical and midlatitude regions are quite similar, and there are differences in the polar regions and stratospheric regions where the AIRS instrument has reduced sensitivity. The atmospheric and surface trends are shown in Table 1, divided into "all" (which is the entire $\pm$ 90 latitude range and 0-1000 mb vertical range) and "T/M" which is the tropical/midlatitude region, which is further reduced to 050-900 mb for air temperature and 300-800 mb for water vapor. "ERA5 direct" are trends computed directly from the geophysical fields, while "ERA5 spectral" are retrieved from the spectral trends.

6.5 Surface emissivity changes

Equation 3 explicitly includes the surface emissivity in the equation of radiative transfer; however Equation 4 assumes this is unchanging. Here we rewrite Equation 4 as

$$\frac{d\overline{BT(\nu)}}{dt} - K_{emissivity}(\nu)\frac{d}{dt}\overline{\epsilon(t)} \rightarrow \frac{d\overline{BT\prime(\nu)}}{dt} = K(\nu)\frac{d}{dt}\overline{X(t)} \tag{6}$$

Ocean emissivity has a dependence on windspeed (Masuda et al., 1988). (Lin & Oey, 2020) and other literature suggest wind speed increases of +2.5 cm s^{-1} yr^{-1} have occured between 1993-2015 in the tropical Pacific, and smaller (or close to zero) values

	dTz/dt K yr^{-1} A SFC-TOA	dTz/dt K yr^{-1} T/M 050-900 mb	dSKT/dt K yr^{-1} A	dSKT/dt K yr^{-1} T/M	dfracWV/dt K yr^{-1} A GND-TOA	dfracWV/dt K yr^{-1} T/M 300-800 mb
ERA5 direct	0.010 ± 0.038	0.029 ± 0.013	0.020 ± 0.035	0.018 ± 0.032	0.003 ± 0.002	0.002 ± 0.001
ERA5 spectral	0.004 ± 0.033	0.027 ± 0.012	0.019 ± 0.033	0.016 ± 0.029	0.001 ± 0.001	0.002 ± 0.001

Table 1. Cosine weighted air temperature, skin temperature, fractional water vapor trends, together with uncertainties. The "ERA5 direct" are directly from the ERA5 geophysical trends, while "ERA5 spectral" are trends retrieved from the converted ERA5 spectral trends.

elsewhere. The monthly ERA5 *u10,v10* 10 m speeds for the 20 year time period in this paper also showed the maximum absolute trend was 0.09 m/s/year (over the Southern Ocean) while the global ocean mean and standard deviation were 0.006 ± 0.022 m s^{-1} yr^{-1}; The emissivity changes over ocean using a 0.025 m s^{-1} wind speed change are on average on the order of 1×10^{-6} per year in the thermal infrared window (or about 0.0003 K yr^{-1} change in the window region); assuming the optical properties of water do not substantially change with the $\sim$ 0.02 K increases seen in all the datasets considered in this paper, these very small emissivity changes are of no consequence.

Land emissivity changes were estimated as follows. A global monthly mean emissivity database, the Combined ASTER and MODIS Emissivity over Land (CAMEL v003) has recently been released (Borbas et al., 2018). We matched the tile centers to the database for the 20 × 12 months spanning our 2002/09 - 2022/08 time period, and computed the emissivity trends over land; the results (not shown here) were on the order of -1×10^{-4} and $+3\times10^{-4}$ in the 800-960 cm^{-1} and 1100-1250 cm^{-1} regions respectively, averaged over the land observations. For each tile the $K_{emissivity}(\nu)\frac{d}{dt}\epsilon(t)$ term was estimated by running SARTA with the default emissivity, then differencing with the SARTA output obtained when the emissivity trends were added on. Averaged over the planet, the spectral changes arising from these emissivity changes were much smaller than the spectral trends seen in Figure 3, about -0.001 K yr^{-1} between 800-960 cm^{-1} and about +0.002 K yr^{-1} on the 1100-1250 cm^{-1} region (which we do not use in our retrieval, since many of the channels are synthetic and the real channels are drifting Strow et al. (2021)). The land only results were roughly three times these magnitudes. Using these emissivity jacobians on the left hand side of Equation 6 and running the retrieval on the adjusted spectral trends over land, resulted in about at most 0.01 K increases to the zonally averaged surface temperature changes over land; zonally averaged these largest differences were at about 40°N to 60°N and -25°S to +15°N, due to emissivity decreases; the 20°N to +35°N region which included the Sahara and swathes of Asia, had emissivity increases but the averaged-over-land temperature decreases were small, as there were offsetting emissivity increases in other land areas at the same latitudes. We did not pursue the impact of these emissivity changes further as the CAMEL database is affected by the stability of the MODIS data, and our results below will not include accounting for changes in land emissivity.

7 Results

The trends retrieved in the previous section using simulated radiance trends show that the retrieval package is working as expected. Here we apply our retrieval to observed AIRS L1C radiance trends and compare the retrieved AIRS_RT geophysical trends to those computed directly from the ERA5/MERRA2 model fields and AIRS L3/CLIMCAPS L3 products. We will have an expectation that since the simulated radiance trends had no noise added to them, the uncertainty in the spectral rates was lower than the actual observed spectral uncertainty; this will lead to larger uncertainties and/or errors in our retrieval using observed radiance trends.

We will make most comparisons against NWP models and L3 products in the context of averages over the descending/night (N) and ascending/day (D) data since the MERRA2 (and GISS) datasets are only available as a D/N average; the reader is referred to the Appendix where we show a few of the D-N differences. The results are shown in the order of surface/column trends (surface temperature and column water), followed by zonal averages of the atmospheric temperature and fractional water vapor trends.

7.1 Skin Temperature trends

There are typically multiple (window) channels that are sensitive to a surface pressure, meaning the radiances typically have more information content for the surface temperature (assuming the surface emissivity is well known and there are no clouds) rather than for example air temperature. Figure 7 shows the diurnally averaged day/night (D/N) surface temperature trends from 6 datasets : AIRS_RT, AIRS L3, CLIMCAPS L3, ERA5, MERRA2 and NASA GISTEMP. AIRS_RT shows an overall global warming of +0.021 K yr^{-1}; the cooling trends include the tropical eastern Pacific and south of Greenland and tropical northern Atlantic. The rest of the datasets also show similar patterns of cooling in the N. Atlantic Ocean, warming over the Arctic and some degree of cooling over the Antarctic Ice Shelf/Southern Ocean as does AIRS_RT. The AIRS v7 L3 shows some cooling over Central Africa and the Amazon not seen in the AIRS_RT trends, where one could expect Deep Convective Clouds and possible cloud clearing issues. We also point out the AIRS L3 product has many missing values off the western coasts of N. and S. America, due to cloud clearing issues. MERRA2 shows more cooling over C. Africa, and just like the AIRS v7 data, a lot of cooling near the Antarctic Ice Shelf. Of note here is that although CLIMCAPS uses MERRA2 as its first guess, their surface temperature trends are not similar, especially around the Antarctic where MERRA2 shows strong cooling trends.Over the ocean GISS shows similar trends to what AIRS_RT trends show. An earlier study of Land Surface Temperatures between 2003-2017 using MODIS (Prakash & Norouzi, 2020) shows very similar large daytime cooling trends over parts of central and western Indian subcontinent that we see from our retrieval as well as directly from the BT1231 channel trends; for tiles that straddle both ocean and land the quantile method picks up the hottest observations, which especially during summer are mostly over the Indian subcontinent. For these reasons we also have confidence in our retrieved cooling trends over for example daytime continental Central/Eastern Africa, which are different from the other four day/night datasets.

The spatial correlations between AIRS_RT retrieved rates and the various datasets is shown in Table 2 while the cosine weighted skin temperature trends are shown in Table 3. By adding in the uncertainty in the trends for any of the individual models or datasets, and then doing the cosine weighting, we estimate uncertainties of about ± 0.015 K yr^{-1} for “ALL”; the uncertainties for “OCEAN” are typically about 2/3 of that value, and for “LAND” are about 4/3 of that value. We emphasize here that we use center point NWP and L3 model data when computing their trends for any grid box, while the AIRS_RT uses the hottest 10% of “clear” data; (Strow & DeSouza-Machado, 2020) showed that the tropical retrieved surface temperature trends and anomalies over ocean correlated very well with those from the ERA-I Sea Surface Temperature dataset.

A notable outlier in this group is the MERRA2 trends, especially over land and the Southern Ocean which are noticeably negative (blue) compared to the other datasets; the agreement with tropical and mid-latitude oceans is much better. As noted earlier,

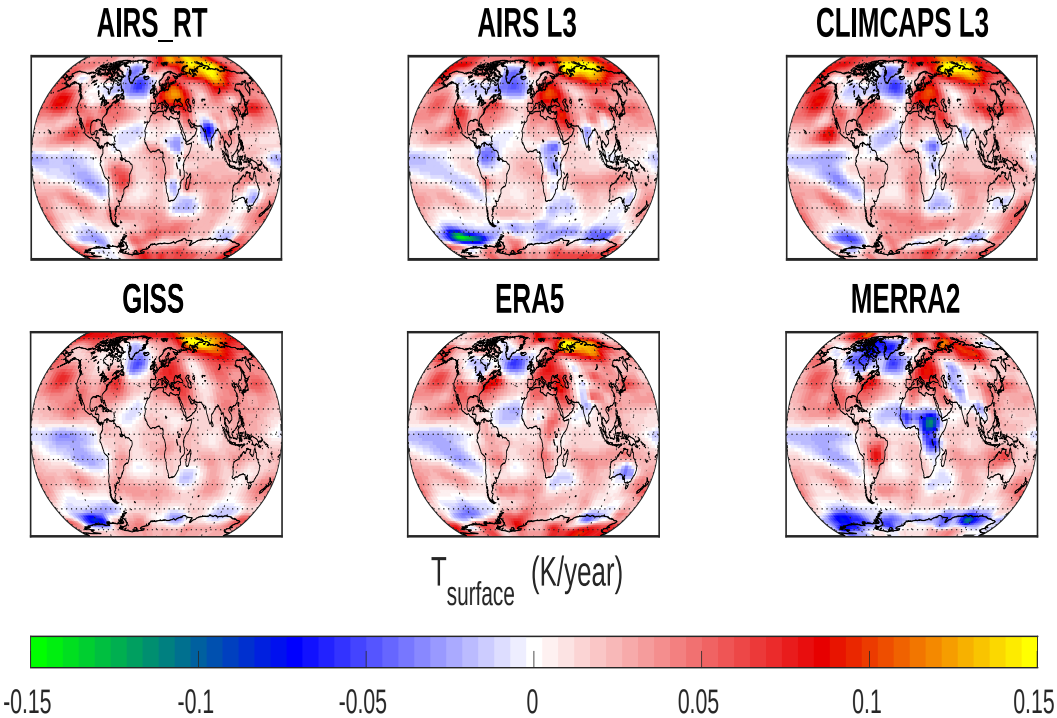

Figure 7. Surface temperature trends dSKT/dt averaged over day and night for AIRS_RT, and from separately fitting the monthly data in ERA5, MERRA2, AIRS L3, CLIMCAPS L3 and GISS.

ERA5	MERRA2	AIRSL3	CLIMCAPSL3	GISS
0.72	0.59	0.80	0.89	0.77

Table 2. Correlations of average (nighttime,daytime) retrieved skin temperature trends from AIRS_RT, versus trends from models/products

the MERRA2 monthly trends come from a combination day/night dataset that was downloaded, which as seen in Figure 7 consists of trends that are both positive and negative, combining to get a closer-to-zero global weighted trend. In addition MERRA2 is the only one of the six that (a) does not have the extreme +0.15 K yr^{-1} warming in the northern polar region and (b) shows a lot of cooling in the Central African area. Using ERA5 monthly data, we devised a test similar to the one mentioned in Section 5 to determine if the differences between MERRA2 and ERA5 surface temperature trends could be due to the temporal sampling (once for MERRA2 versus eight times for ERA5). For each month we matched the eight ERA5 timesteps available per month to the tile centers and then averaged the surface temperatures per month; the ensuing geophysical timeseries was then trended. The day/night ERA5 average of Figure 7 was compared to these trends; of note are (a) we did not see the cooling in Africa and near the Antarctic that is seen in MERRA2 and (b) the main differences between the 1.30 am/1.30 pm average in the bottom middle (ERA5) panel were over land (all 5 continents); the histograms of the differences showed the peak was typically close to 0 K yr^{-1}, but the widths over land were about ± 0.02K yr^{-1} or less (compared to ± 0.005 K yr^{-1} over ocean). Both AIRS L3 and MERRA2 show cooling in the Southern Ocean; we note that although MERRA2 is the *a-priori* for CLIMCAPS L3, their trends are different that those from MERRA2; in

SKT trend K yr^{-1}	AIRS_RT	AIRS	CLIMCAPS	ERA5	MERRA2	GISS
ALL	0.020	0.017	0.021	0.023	0.011	0.021
TROPICS	0.011	0.011	0.012	0.016	0.010	0.015
MIDLATS	0.029	0.020	0.028	0.026	0.020	0.026
POLAR	0.032	0.028	0.033	0.041	-0.005	0.028
OCEAN	0.019	0.011	0.019	0.017	0.012	0.017
LAND	0.022	0.030	0.024	0.038	0.010	0.030

Table 3. Cosine weighted skin temperature trends; uncertainties are on the order of ± 0.015 K as explained in the text.

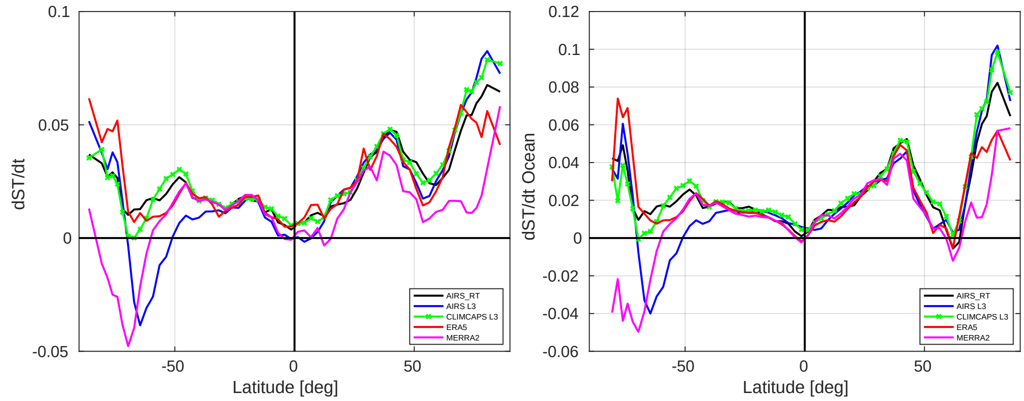

Figure 8. Zonally averaged surface temperature trends for (left) sum of ocean and land point and (right) ocean only.

fact AIRS_RT shows the closest correlation to the observational CLIMCAPS L3 trends. The AIRS L3 trends in the Southern Ocean region could arise because of problems identifying ice during the L2 retrieval (private communication : Evan Manning (JPL) and John Blaisdell (NASA GSFC)) though the MERRA2 trends also show significant cooling in that region, where few surface observations from buoys poleward of 60°exist to help resolve these differences (see for example Figure 10 in (Haiden et al., 2018)).

Figure 8 shows the zonally averaged total (land+ocean) and ocean only surface temperature trends. The equator to midlatitude ocean trends are almost linear for all datasets, with the slope for the northern hemisphere being about double that of the southern hemisphere (roughly 0.001 K yr^{-1} per deg latitude). Again focusing on the right hand plot, the AIRS L3 trends are negative in the Southern Ocean regions, compared to the other 3 datasets, due to the cooling trends around the Antartic continent shown earlier, but then agrees with most of the other datasets over the Antartic; the MERRA2 trends significantly differ between -90 S and -50 S. MERRA2 and ERA5 also show slightly smaller warming trends in the Northern Polar, compared to the three AIRS-based datasets.

We point out that the trends seen in Figure 7 vary noticeably at more local, regional levels and furthermore this spatial variation can differ between daytime and nighttime, evident in Figure A1 of Appendix Appendix A, and that the observational sets (AIRS_RT,

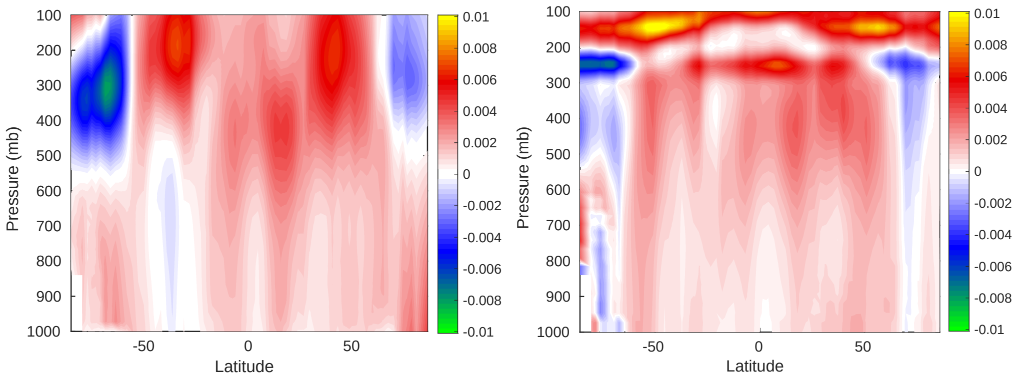

Figure 9. dWVfrac/dt (left) without and (right) with MLS *a-priori* in the upper atmosphere

CLIMCAPS L3 and AIRS L3) had larger differences than ERA5. Discussing the possible causes is outside the scope of the paper.

7.2 Addition of Microwave Limb Sounder Water Vapor A-priori

The Microwave Limb Sounder (MLS), on board NASA's Aura platform, is designed for sounding of the atmosphere above 300 mb. We computed water vapor trends from the L3 data produced for that instrument (above 300 mb) and used them as an *a-priori* for the AIRS_RT retrieval.

Figure 9 shows the retrieved fractional water vapor trends when the *a-priori* trend in the upper atmosphere in the left and right panels were zero, or used MLS trends, respectively. One sees that the additional information brought in by the instrument sensitive to upper troposphere humidity, significantly changes the water vapor sounding especially in the polar region by moving towards the MERRA2 and ERA5 fractional water vapor trends seen in Figure 12. We note that the other related results shown in this paper use the MLS *a-priori*.

7.3 Column water vapor trends

Column water is dominated by water vapor amounts close to the surface and the column vapor trends thus provide an assessment of the water vapor retrieval quality in the lower atmosphere. The water vapor information in the lowest layers is best retrieved using the weak water lines in thermal infrared region. As noted earlier this part of the retrieval is significantly complicated by the simultaneous presence of nonzero surface temperature, air temperature and water vapor jacobians in this spectral region, meaning the AIRS instrument has much reduced sensitivity to the water vapor amounts in these lowest layers. In addition the changing concentration of very minor gases such as CFC-11 and CFC-12 (Strow & DeSouza-Machado, 2020) are quite evident in the spectral trends, further complicating the water vapor trend retrieval for the lowest layers.

Figure 10 shows the zonally averaged column water vapor trends; not shown are the error bars which are on the order of ± 0.005 mm/year. AIRS_RT is from our retrievals while the rest are directly from the NWP or L3 model fields. Close examination

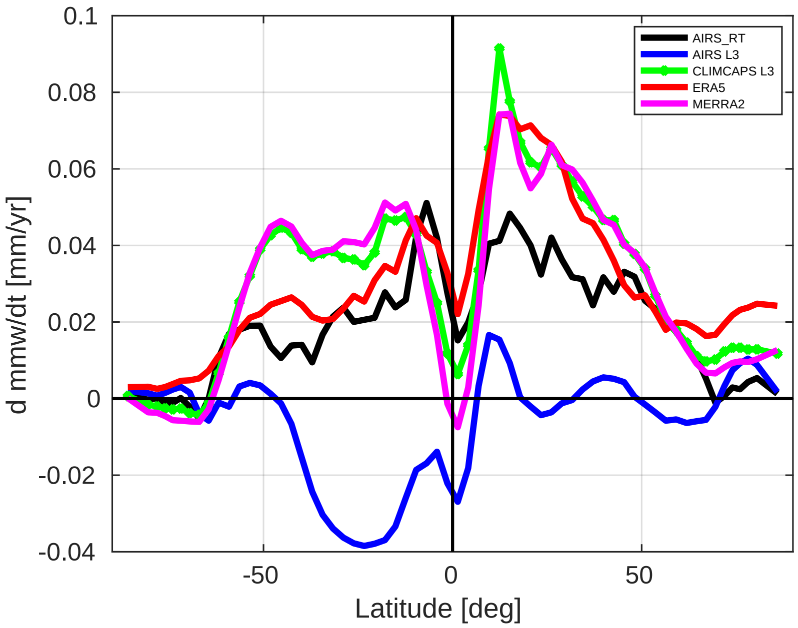

Figure 10. Zonally averaged column water vapor trends for AIRS_RT, AIRS L3, CLIMCAPS L3, ERA5 and MERRA2.

shows the CLIMCAPS L3 column water trend is nearly identical to the MERRA2 trend, as is also seen in lower atmosphere water vapor trends shown later in Figure 12. Conversely the column water vapor trends for AIRS L3 are negative in the lower troposphere in the midlatitudes and tropics, which is not to be expected given that the surface temperature trends are positive. AIRS_RT nominally agrees with ERA5 and MERRA2 in the tropics and midlatitudes, but is smaller than either in the northern polar regions. A reduced rate for AIRS_RT is additionally seen in the 0-50 N latitudes, where there is a larger fraction of land (for which we do not use the assumption of constant relative humidity) compared to the Southern Hemisphere. Screening out the tiles over land slightly improves the agreement between reanalysis (ERA5, MERRA2) vs AIRS_RT column water trends. Examination of the spectral trends in the window region does not shed any more insight into the differences, as the observation spectral trends and NWP reconstructed trends are very similar and we are fitting the observed trends. The magnitudes and patterns look similar to the 2005-2021 column water trends shown in (Borger et al., 2022), which were derived using observations from the Ozone Monitoring Instrument (OMI). We point out their 16 year zonally averaged trends look similar to the 20 year ERA5 zonally averaged column water trends between -60°S and -10°S, but become almost a factor of 2 larger between -10°S and +40°N; the zonally averaged OMI 16 year trends are negative in the polar regions. The column water trends are summarized in Table 4.

D/N differences (not shown) for AIRS_RT were on the order of $\pm$ 0.005 mm yr^{-1} (with daytime trends being smaller over land), for AIRS L3 were on the order of $\pm$ 0.01 mm yr^{-1} or more (with larger values happening over the daytime tropical oceans), while that for ERA5 and CLIMCAPS L3 were typically on the order of $\pm$ 0.03 mm yr^{-1} or less.

7.4 Zonal atmospheric temperature and water vapor trends

DATASET mm yr^{-1}	OMI 16 years	AIRS_RT 20 years	ERA5 20 years	MERRA2 20 years	AIRS L3 20 years	CLIMCAPS L3 20 years
GLOBAL (cosine average)	0.051	0.021	0.035	0.036	-0.009	0.038
TROPICAL	0.083	0.028	0.047	0.042	-0.015	0.045

Table 4. Column water trends based on OMI data (16 years) and AIRS_RT, ERA5 and MERRA2 (20 years). The units are in mm yr^{-1}; the uncertainties are on the order of 0.1 mm yr^{-1} for OMI and AIRS_RT, and half that for ERA5 and MERRA2, and AIRS L3 and CLIMCAPS L3.

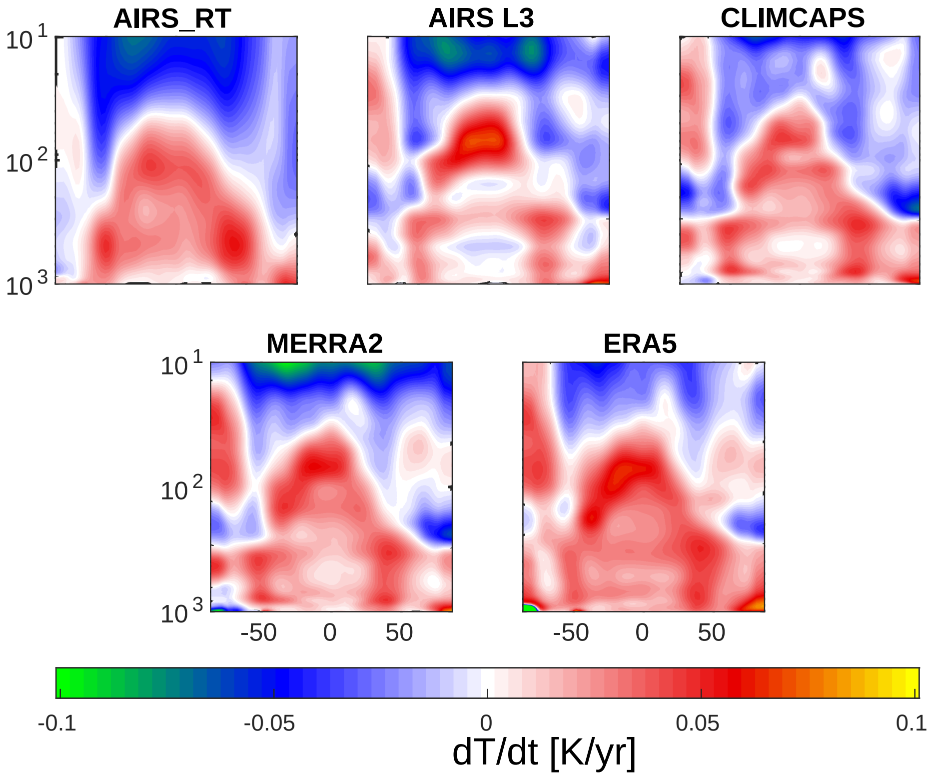

Figure 11. Zonally averaged dT/dt shown in 5 panels. Horizontal axis is latitude while vertical axis is pressure. The y-limits are between 10 to 1000 mb, on a logarithmic scale.

Figure 11 shows the zonally averaged atmospheric temperature trends from five of the datasets in Figures 7,10 above. In the troposphere the AIRS_RT retrievals show the same general features as the trends from ERA5, though they begin to diverge in the stratosphere and especially above that. In particular AIRS_RT does not show warming in the Southern Polar stratosphere; we have separately looked into seasonal trends and noted that our retrieved September/October/November temperature trends in the upper atmospheric Southern Polar regions are on the order of -0.12K yr^{-1}, possibly leading to an overall no net heating/cooling for the annual trends. In addition we point out that both our results and AIRS v7 L3 show a hint of cooling over the tropical surfaces. Note that CLIMCAPS is initialized by MERRA2, and their temperature trends are quite similar. AIRS v7 looks similar to AIRS_RT except in the tropics where it almost has cooling in the lower troposphere and much more warming in the lower stratosphere. The correlations between AIRS_RT and the [AIRS L3, CLIMCAPS L3, MERRA2, ERA5] temperature trends of Figure 11 are [0.74,0.65,0.74,0.72] respectively.

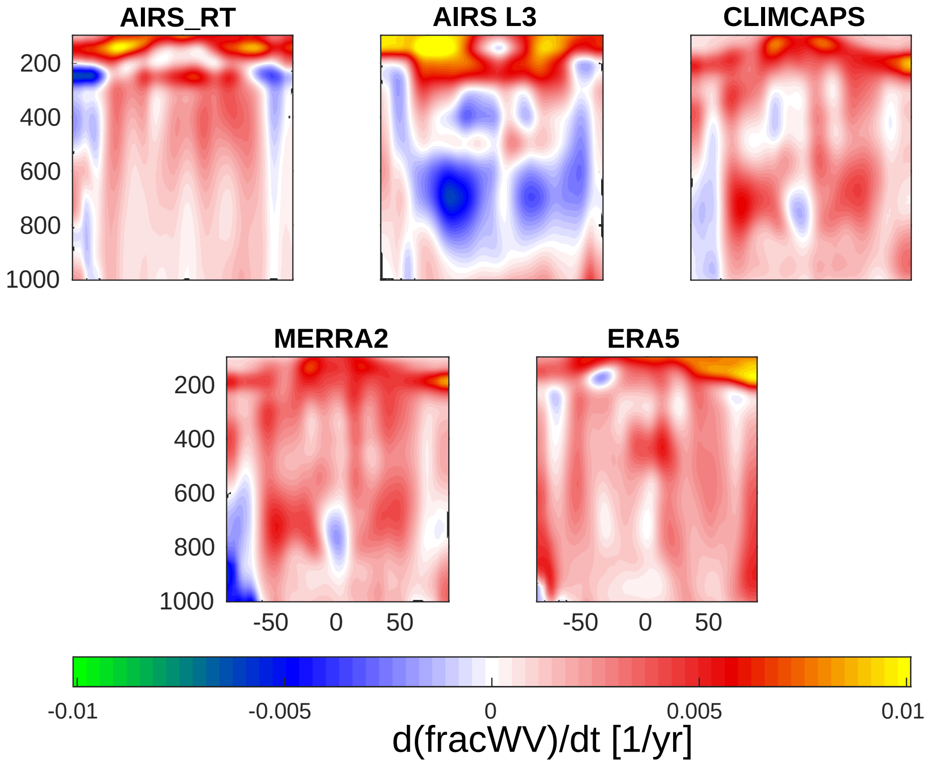

Figure 12. Zonally averaged dWVfrac/dt shown in 5 panels. Horizontal axis is latitude while vertical axis is pressure. The y-limits are between 100 to 1000 mb, on a linear scale.

Figure 12 shows the zonally averaged atmospheric fractional water vapor trends (d/dt WV(z,t)/<WV(z,t)>). The five panels are markedly different from one another. The AIRS_RT trends resemble those of ERA5 in the tropical troposphere, though we do not have drying in the lower tropical layers. Conversely, the observed trends in the Southern Polar (AIRS L3, CLIMCAPS L3 and AIRS_RT) show drying rather than wetting, though AIRS_RT is less than that of CLIMCAPS/MERRA2. AIRS_RT is an outlier in the upper polar atmosphere trends, as both the signals and the jacobians are close to zero. Of some concern is a little bit of drying in the northern polar region, where there are low H_2O amounts leading to small jacobians. CLIMCAPS v2 looks quite similar to the MERRA2 trends. AIRSv7 shows substantial drying in the lower troposphere, and considerable wetting in the upper troposphere, compared to any of the other datasets. Spectral closure studies (using the AIRS v7 H_2O trend $\times$ the H_2O jacobians derived above from ERA5 average profiles) are not shown here, but differ noticeably from the CCR trends from AIRS v7 in the 1300-1600 cm^{-1} region, indicating there are inadequacies in the AIRS V7 water vapor retrievals. The correlations between AIRS_RT and the [AIRS L3, CLIMCAPS L3, MERRA2, ERA5] fractional water vapor trends of Figure 12 (limited to 100 mb, 1000 mb) are [0.65,0.24,0.36,0.58] respectively.

Figure 13 shows the 400 mb fractional water vapor trends, with the left panel being the AIRS_RT trends while the right panel is the ERA5 trends. Note that there is general agreement except in the Southern Polar region, as also seen later in Figure 12 in the other two observational L3 datasets (AIRS v3 and CLIMCAPS). This could be related to work by (Boisvert et al., 2019) who showed decreasing evaporation from the Southern Ocean in the 2003-2016 period due to increasing ice cover.

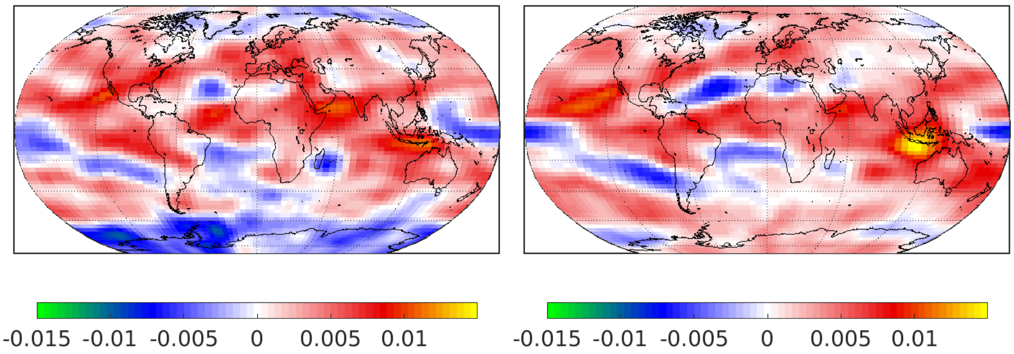

Figure 13. The 400 mb fractional water vapor trends for (left) AIRS_RT and (right) ERA5 show general agreement except in the Southern Polar Regions.

8 Uncertainty

The uncertainties for the AIRS v7 geophysical products are impacted by radiance noise amplification due to cloud clearing (Susskind et al., 2003) and the neural net first guess, while state vector errors are estimated based on regressions. CLIMCAPS L2 geophysical products are similarly impacted by cloud clearing noise in the radiances, but these are fully propagated together with geophysical error estimates from the MERRA2 first guess, through the retrieval algorithm which uses Optimal Estimation (Smith & Barnet, 2020). No estimate of uncertainties are available for the monthly L3 products.

The uncertainties for the AIRS_RT trends is much more straightforward : the spectral uncertainties shown in Figure 4 are used together with the state vector covariance matrices to generate the uncertainty matrix using the relevant equations of Optimal Estimation Rodgers (2000); we use the diagonal elements for the final uncertainties. Panels (A) and (C) of Figure 14 shows the zonally averaged (D/N) uncertainties as a function of pressure and latitude. Inspection of the radiance trends uncertainties shown in the center panel of Figure 4 shows the upper atmosphere temperature sounding region (650-700 cm^{-1}) has much larger uncertainty in the polar regions. The instrument and spectroscopy characteristics, coupled with these observational uncertainties, are such that for temperature the smallest errors are in the tropics while the largest errors are in polar upper atmosphere, which are the regions below 100 mb where the ERA5 trends differ most from AIRS_RT trends. Similarly for water vapor the larger errors are in the lower atmosphere and above about 300 mb; the constant RH assumption and MLS *a-priori* help alleviate the errors in the retrieved trends.

The *Z-test* confirmed this picture, as seen in panels (B) and (D) of Figure 14, which show the temperature and fractional water vapor trends, together with black dots marking the (latitude,altitude) points where the trends are larger than the uncertainty in the trends, at the 5% significance level. This happens in panel (B) for the temperature trends in most of the tropical/mid-latitude free troposphere (and stratosphere) but not at the southern polar stratosphere; and in panel (D) for fractional water vapor trends in the 200-600 mb range, from the Southern Polar region to about +60 N latitude, and some spots in the Northern Polar.

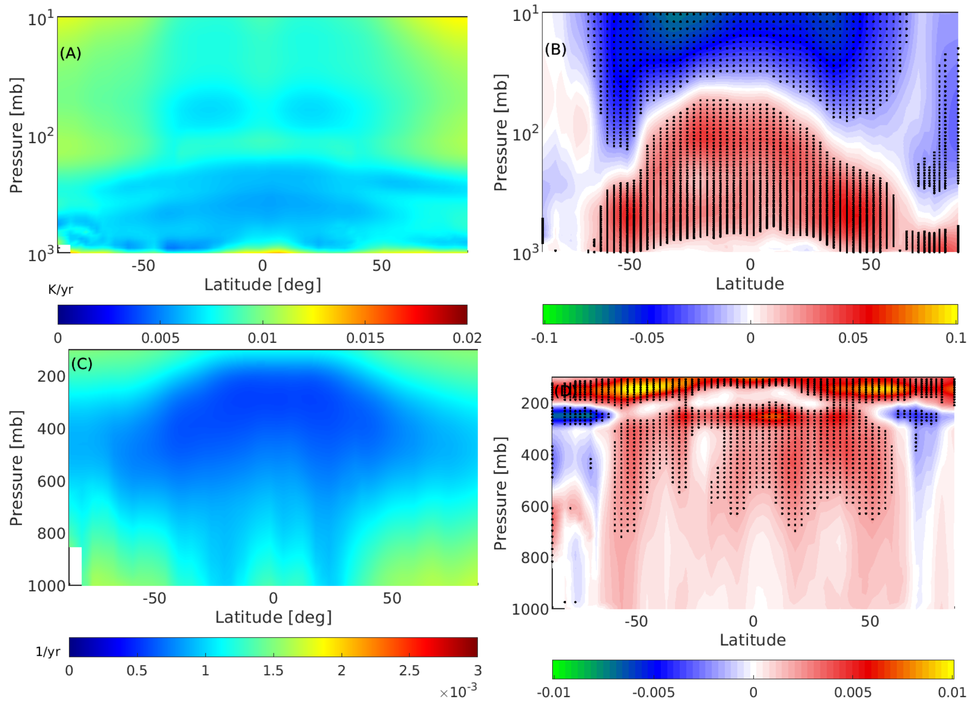

Figure 14. Zonally averaged D/N plots of (A) temperature uncertainties in K yr^{-1} and (B) temperature trends in K yr^{-1} together with null hypothesis. (C) and (D) are the same except for fractional water vapor uncertainty and trends in 1/year. See text for more detailed explanation.

9 Discussion

In general for surface temperature trends, the disagreements between the six sets shown in Figure 7 are over the polar regions and over land (especially over the Amazon and Central Africa) and are smallest over tropical and mid-latitude oceans, indicating the best agreements, except for slightly larger differences off the western coast of the Americas and Africa (which have a prevalence of MBL clouds). The atmospheric temperature trends in general agreed except for the upper atmosphere polar regions and in the high altitudes (less than about 200 mb). Similarly fractional water vapor trends differed most in the upper atmosphere (200 mb and above) and in the tropical/mid-latitude 600-800 mb region. A quick glance at Figure 12 shows the former is due to lower sensitivity to upper atmosphere water vapor, leading the AIRS_RT retrievals to have low values while the AIRS L2 retrieval is initialized by a neural net; conversely the latter is due to the AIRS L3 retrieval being negative while the rest were mainly positive. Similarly the AIRS_RT retrieval differs above the Antarctic continent.

In general the observed surface temperature trends from the AIRS_RT retrievals agree with the ERA5 and MERRA2 trends, as well as the NASA GISS trends, except in the Southern Antarctic. That is a region where there are few surface observations; for retrievals there are competing effects of using ice vs ocean surface emissivity. Overall, the AIRS_RT retrieved surface temperature trends are typically in between ERA5 and MERRA2 for land + ocean in all regimes (tropical, midlatitude and polar), though slightly larger overall for ocean than the two reanalysis datasets; in general they are closer to the ERA5 trends than the MERRA2 trends.

(Strow et al., 2021) demonstrated that the long- and medium- wave channels of the AIRS instrument are radiometrically stable to better than 0.002-0.003 K yr^{-1}, which is much smaller than the surface and tropospheric temperature trends in the reanalysis models, AIRS L3 data and our retrieved trends. A separate analysis of spectral trend uncertainties after 05,10,15,20 years (not shown here) show that these uncertainties have been steadily decreasing and are now approaching this number, as can be seen in the bottom left panel of Figure 3. Furthermore, though we cannot guarantee only cloud free scenes in our chosen Q0.90 dataset used in this paper, the high correlations between other dataset surface trends compared to ours, is a good indication that our results come from mostly cloud-free scenes, or scenes whose clouds have negligible impact on our results.

The observed zonal temperature trends agree with those from the models and the AIRS L3 products, except in the polar regions. Again this could be an issue of using slightly incorrect surface emissivity for the AIRS_RT retrievals. In addition we point out that since there is very little water vapor, the temperature jacobians near the surface are quite small in magnitude (compared to more humid atmospheres) and so it is difficult to separate out the effects of surface temperature trends versus lower atmosphere temperature and H_2O trends. The quantile construction used in this paper means that for example tiles straddling the subcontinent of India and the ocean will preferentially pick the land surface observations for daytime, which could lead to misleading trends on these coastal tiles. It is possible to subdivide the $3° \times 5°$tiles into for example $1° \times 1°$grids and do the analysis, but the number of observations per small grid cell would drop, leading to more noise in the retrieved trend.

In general the AIRS_RT retrieved absolute column water trends are slightly smaller than ERA5 in the Southern Hemisphere but noticeably smaller in the Northern Tropics to midlatitudes. We have mentioned difficulties we have retrieving H_2O close to the surface and in the upper atmosphere, due to the known sensitivity of infrared sounders whose water vapor averaging kernels peak in the 300-600 mb range, and we have pointed out examination of the spectral residuals in the window region shows we are fitting the signal. The MERRA2 and CLIMCAPS column water vapor trends are quite similar, while the AIRSv3 L3 trends are noticeably different, being negative almost everywhere.

Given the complex numerical algorithms used in both the reanalysis models and the AIRS L3 retrievals as well as those in the AIRS_RT trends, it is difficult to offer precise explanations for any of the trends shown above. Our results are relatively robust to changes in the covariance or Tikonov parameter settings. For instance changing them by factors of two would keep the trends about the same, though of course the uncertainties would change. There are however a few general points that can be made. The first is that since infrared instruments are sensitive to the 300-800 mb region and lose sensitivity outside this, the retrievals from AIRS_RT and AIRS L3 have difficulties with water vapor in the lower (Planetary Boundary Layer) and upper troposphere/lower stratosphere. One way to mitigate this is to use trended data from external sources in the *a-priori*, while keeping the *a-priori* trends for all other parameters as 0. For example we have shown we can use the MLS data above 300 mb without significantly degrading the AIRS_RT retrieval in the middle and lower atmosphere; conversely the CLIMCAPS retrievals are initialized by MERRA2 and while they can pull out weather signals, their L3 trends are still quite closely tied to the MERRA2 trends. The tropical and mid-latitude ocean surface temperature trends from the numerical models that assimilate data, L3 products and AIRS_RT are very similar; however they start to show differences where there are few *in-situ* data combined with problems with ice identification (surface emissivity)/cold temperatures which exacerbate the drifting AIRS detector problems (Strow et al., 2021), such as the Arctic and Southern Ocean.

10 Conclusions

We have described a novel method to obtain global thermodynamic atmospheric climate trends, starting from infrared allsky hyperspectral observations which are then subset for "nominally clear" scenes. Our retrieved trends are derived using trends from well characterized (radiometrically stable) radiances and from zero *a-priori* (except for a constant relative humidity assumption). This makes them much more direct and traceable than trends from traditional L2 retrieval algorithms, which use complicated *a-priori* information. We also did "radiative closure" tests by running the monthly NWP or L3 fields through a radiative transfer model to compare the spectral trends so obtained against the observed spectral trends, which showed the most disagreement in the water vapor sounding regions.

The temperature and water vapor trends retrieved from the "nominally clear" radiance trends resemble those computed from monthly ERA5 and MERRA2 reanalysis. The radiative spectral closure helps identify the cause of differences in the geophysical trends, rather than solely attributing them to deficiencies (eg the well known reduced sensitivity to water vapor near the boundary layer and above 200 mb) with our retrieval. For example the AIRS_RT temperature trends are quite similar to the reanalysis (MERRA2/ERA5) trends, while the water vapor (and/or Relative Humidity) trends are quite different, especially in the lower troposphere and upper troposphere, which is clearly manifest as differences in the spectral trends in the water vapor sounding region.

The 20 years of AIRS observations were binned into nominal 3×5 degree grid boxes covering the planet, with a time step of 16 days, from which anomalies and trends were obtained. To alleviate the reduced sensitivity of hyperspectral sounders to water vapor in the lower atmosphere we used an assumption of 0.01 increase in relative humidity to initialize the *a-priori* lower atmosphere fractional water vapor rates, while we similarly used Microwave Limb Sounder trends as an *a-priori* to address the high altitude water vapor deficiencies caused by lower sensitivity to upper atmosphere water vapor. New or updated time dependent surface emissivity databases may become available in the future, enabling us to include those effects into Equation 4. Problems in the polar regions and Planetary Boundary Layer water vapor retrievals will be harder to overcome since there is very little sensitivity to water vapor in these regions, together with fewer observations to compare against, though more work is planned to address both of these.

In this paper we used the 90th quantile (Q0.90) nominally "hottest" observed BT1231 data to form a time series over which to obtain radiance trends, after establishing that the spectral trends from this quantile differed by less than about $\pm$ 0.0015 K yr^{-1} from the 50th (or average) quantile. In the future we plan to base the data subset selection on MODIS cloud products (obtained at 1 km resolution compared to the AIRS 15 km resolution). In any case the AIRS L1C Q0.90 spectral trends used for the AIRS_RT results are very comparable to trends from quality assured binned AIRS CCR data (Manning, 2022). The quantile method allows us to select which data to use in the trends : we have explored doing the trend retrievals using the cloud fields contained in ERA5, together with the TwoSlab cloud algorithm (De Souza-Machado et al., 2018) to compute jacobians when clouds are present, together with trends from the Q0.50 dataset described above. The retrieved geophysical trends resemble those described above in the mid to upper atmosphere, and differ in the lower atmosphere, but more work is needed on this and is not discussed further. Longwave clear sky flux trends (both outgoing top-of-atmosphere and incoming bottom-of-atmosphere) and climate feedbacks will be discussed in a separate paper.

While the Aqua platform is scheduled to be terminated within the next few years, copies of near identical CrIS instruments are already in orbit, and more will be launched over the next few years, till at least 2040. The Climate Hyperspectral Infrared Radiance Product (CHIRP) (Strow et al., 2021) will seamlessly combine the AIRS data between 2002-2015 to CrIS data from 2015-2040 to obtain a 40 year observational radiance record over which to study climate. This availability means that AIRS_RT and future AIRS/CrIS versions, is well positioned to enable climate analysis of geophysical trends for years to come.

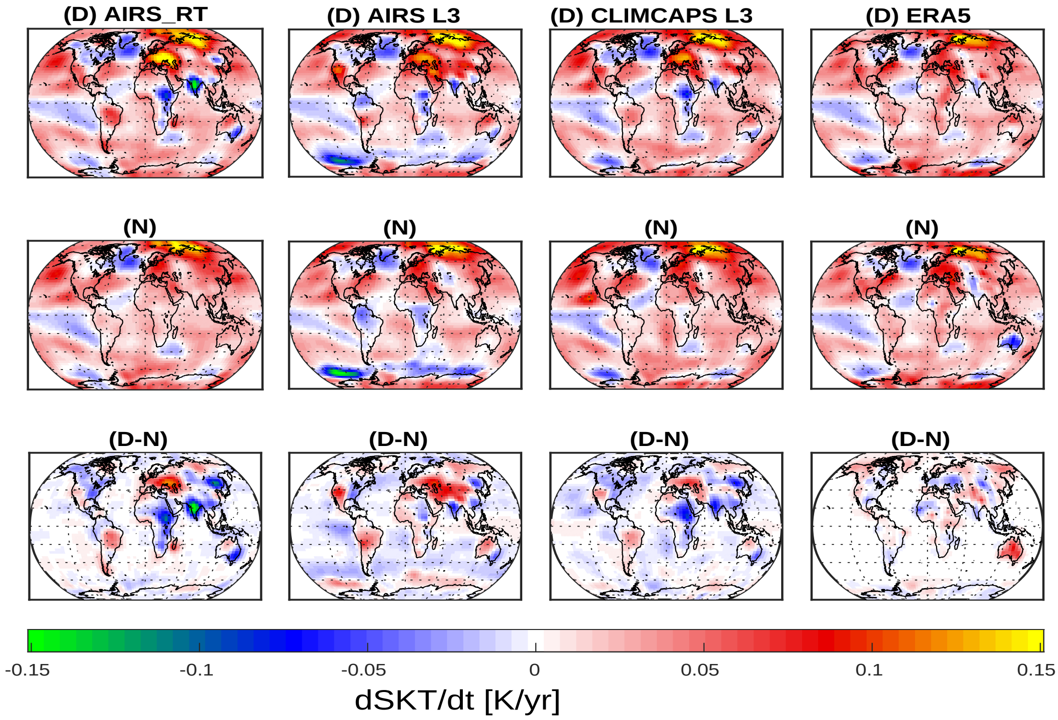

Figure A1. Top two rows : The (top) day and (middle) night surface temperature trends for AIRS_RT, AIRS L3, CLIMCAPS L3 and ERA5. Third row (bottom) is the D-N difference.

Appendix A Day versus Night surface temperature trend differences

Figure A1 shows the (top) daytime and (middle) nighttime surface temperature trends; from left to right the datasets are (observational) AIRS_RT, AIRS L3, CLIMCAPS L3 and (reanalysis) ERA5. In general the AIRS observational datasets show enhanced daytime cooling over the Indian subcontinent and Central Africa, compared to the ERA5 model; they also show daytime warming trends over continental Europe and central Asia and the Amazon are larger than during the nightime. With the large ocean heat capacity and smaller land heat capacity, the land is expected to show more of a diurnal cycle than ocean. ERA5 sees warming over Eastern/Central Africa during daytime while the observational datasets see cooling. Similarly the three observational datasets see more daytime cooling over the Indian sub-continent and south eastern Australia than does ERA5; we omit more detailed analysis in this paper. During the nighttime, the AIRS L3 product has cooling over C. Africa and parts of the Amazon. The day-night differences are seen in the bottom row of the same figure. Note the colorbar is the same for all three rows. The differences are close to zero over the ocean. AIRS_RT and CLIMCAPS L3 see more daytime cooling over E. Africa and the Indian subcontinent. Overall the magnitude of the day - night differences for the observations are larger for the AIRS observational datasets than for ERA5. ERA5 also sees negative differences over Central Asia compared to the AIRS observational datasets, which see positive differences (higher surface temperature trends during the daytime).

The atmospheric temperature and fractional water vapor day-night differences are quite small (compared to the average values) and not shown here; AIRS L3 shows noticeable more wetting of the 600-800 mb region during daytime versus nightime, compared to the other three.

Data Availability Statement

The observation data used in this paper (AIRS L1C radiances), as well as the AIRS L3, CLIMCAPS L3, MERRA-2 and Microwave Limb Sounder monthly data products, are freely available to the public on the NASA Goddard Space Flight Center Earth Sciences (GES) Data and Information Services Center (DISC) servers https://disc.gsfc.nasa.gov/. Monthly ERA5 is freely available through (single levels) https://cds.climate.copernicus.eu/datasets/reanalysis-era5-single-levels-monthly-means?tab=overview and (pressure levels) https://cds.climate.copernicus.eu/datasets/reanalysis-era5-pressure-levels-monthly-means?tab=overview. GISTEMP monthly model output are also freely available from https://data.giss.nasa.gov/gistemp/. The Matlab based source code used for the analysis is freely available on https://github.com/sergio66/oem_climate_code, while the F90 kCARTA (De Souza-Machado et al., 2018) line-by-line code used to make the jacobians is freely available on https://github.com/sergio66/kcarta_gen.

Acknowledgments

We gratefully acknowledge the help of H. Motteler, who designed, implemented and ran the AIRS tiling code. C. Hepplewhite helped test the outputs of the tiling code, S. Buczkowski worked on bringing down almost all the data used in this study, and made the uniform/clear datasets. Chris Barnet has given helpful comments and general advice regarding retrievals and comparisons to the AIRSv7 and CLIMCAPS datasets. The hardware used is part of the UMBC High Performance Computing Facility (HPCF). The facility is supported by the U.S. National Science Foundation through the MRI program (grant nos. CNS–0821258, CNS–1228778, OAC–1726023, and CNS—1920079) and the SCREMS program (grant no. DMS–0821311), with additional substantial support from the University of Maryland, Baltimore County (UMBC). See hpcf.umbc.edu for more information on HPCF and the projects using its resources.

References

2023, G. T. (2005). *GISS Surface Temperature Analysis (GISTEMP), version 4* (Tech. Rep.). data accesed 2023/11/30 at https://data.giss.nasa.gov/gistemp/: NASA Goddard Institute for Space Studies.

Boisvert, L., Vihma, T., & Shie, C.-L. (2019). Evaporation From the Southern Ocean Estimated on the Basis of AIRS Satellite Data. *JGR*, *125*, https://doi.org/10.1029/2019JD030845.

Borbas, E., Hulley, G., Feltz, M., Knuteson, R., & Hook, S. (2018). The Combined ASTER MODIS Emissivity over Land (CAMEL) Part 1: Methodology and High Spectral Resolution Application. *Remote Sens.*, *10*, 643; https://doi.org/10.3390/rs10040643.

Borger, C., Beirle, S., & Wagner, T. (2022). : Analysis of global trends of total column water vapour from multiple years of OMI observations. *Atmos. Chem. Phys*, *22*, 10603–10621, https://doi.org/10.5194/acp-22-10603-2022.

Clough, S., Shephard, M., Mlawer, E., Delamere, J., Iacono, M. J., Cady-Pereira, K., ... Brown, P. (2005). Atmospheric radiative transfer modeling : a summary of the AER codes. *J. Quant. Spectrosc. Rad. Trans.*, *91*, 233-244doi:10.016/j.qsrt2004.05.058.

De Souza-Machado, S., Strow, L. L., Motteler, H., & Hannon, S. (2020). kCARTA : A Fast Pseudo Line by Line Radiative Transfer Algorithm with Analytic Jacobians, Fluxes, Non-Local Thermodynamic Equilibrium and Scattering. *Atmos. Meas. Tech.*, *31*, 323–339, https://doi.org/10.5194/amt-13-323-2020.

De Souza-Machado, S., Strow, L. L., Tangborn, A., Huang, X., Chen, X., Liu, X., ... Yang, Q. (2018). Single-footprint retrievals for AIRS using a fast TwoSlab cloud-representation model and the SARTA all-sky infrared radiative transfer algorithm. *Atmos. Meas. Tech.*, *11*, 529-550,https://doi.org/10.5194/amt-11-529-2018.

Gelaro, R., & Coauthors. (2017). MERRA-2 Overview: The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2). *J. Clim.*, *30*, 5419-5454, doi: 10.1175/JCLI-D-16-0758.1.

Gordon, I., & Rothman, L. e. a. (2022). The HITRAN 2020 molecular spectroscopic database. *J. Quant. Spectrosc. Rad. Trans.*, *277*, 1-82,https://doi.org/10.1016/j.jqsrt.2021.107949.

Haiden, T., Dahoui, M., Ingleby, B., de Rosnay, P., Prates, C., Kuscu, E., ... Jones, L. (2018). *Use of in situ surface observations at ECMWF* (No. 834).

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., ... Thépaut, J.-N. (2020). The ERA5 Global Reanalysis. *Quart.J.Roy.Meteorol.Soc.*, *146*, 1999-2049,https://doi.org/10.1002/qj.3803.

Huang, X., Chen, X., Fan, X., Kato, S., Loeb, N., Bosilovich, M., ... Strow, L. (2023). A Synopsis of AIRS Global-Mean Clear-Sky Radiance Trends From 2003 to 2020. *JGR*, *127*, e2022JD037598.https://doi.org/10.1029/2022JD037598.

Irion, F., Kahn, B., Schreier, M., Fetzer, E., Fishbein, E., Fu, D., ... Yue, Q. (2018). Single-footprint retrievals of temperature, water vapor and cloud properties from AIRS. *Atmos. Meas. Tech.*, *11*(2), https://doi.org/10.5194/amt-11-971-2018.

King, M., Platnick, S., Menzel, P., Ackerman, S., & Hubanks, P. (2013). Spatial and Temporal Distribution of Clouds Observed by MODIS Onboard the Terra and Aqua Satellite. *IEEE*, *51*, 3826-3852,10.1109/TGRS.2012.2227333.

Lambert, A., Read, W., Livesey, N., & Fuller, R. (2021). *MLS/Aura Level 3 Monthly Binned Vapor (H2O) Mixing Ratio on Assorted Grids V005* (Tech. Rep.). `https://disc.gsfc.nasa.gov/datasets/ML3MBH2O_005/summary`: NASA Greenbelt, MD, USA, Goddard Earth Sciences Data and Information Services Center (GES DISC).

Lenssen, N., Schmidt, G., Hansen, J., Menne, M., Persin, A., Ruedy, R., & Zyss, D. (2019). Improvements in the GISTEMP uncertainty model. *JGR*, *124*(12), 6307-6326, doi:10.1029/2018JD029522.

Lin, Y., & Oey, L. (2020). Global Trends of Sea Surface Gravity Wave, Wind, and Coastal Wave Setup. *J. Clim.*, *33*, 769–785, https://doi.org/10.1175/JCLI-D-19-0347.1.

Maddy, E., & Barnet, C. (2008). Vertical Resolution Estimates in Version 5 of AIRS Operational Retrievals. *IEEE*, *66*(8), 2375-2384, DOI 10.1109/TGRS.2008.917498.

Manning, E. (2022). *Nasa jpl private communication, 4/2022.* (AIRS CCR L3 data)

Manning, E., Aumann, H., Broberg, S., Pagano, T., Wilosn, R., Yanovsky, I., & Strow, L. (2020). *Eos airs l1c data product user guide* (Tech. Rep.). https://docserver.gesdisc.eosdis.nasa.gov/public/project/AIRS/V6.7_L1C_Product_User_Guide.pdf: NASA.

Masuda, K., Takashima, T., & Takayama, Y. (1988). Emissivity of pure and sea waters for the model sea surface in the infrared window regions. *Remote Sensing of Environment*, *24*, 313–329.

Palchetti, L., Brindley, H., Bantges, R., Buehler, S. A., Camy-Peyret, C., Carli, B., ... Serio, C. (2020). FORUM: unique far-infrared satellite observations to better understand how Earth radiates energy to space. *Bull.Amer.Met.Soc.*, *101*, https://doi.org/10.1175/BAMS-D-19-0322.1.

Peters, W., Jacobson, A., Sweeney, C., & Tans, P. (2007). An atmospheric perspective on North American carbon dioxide exchange: CarbonTracker. *Proc. Nat. Aca. Sci.*, *104*, 18925-18930,https://doi.org/10.1073/pnas.0708986104.

Prakash, S., & Norouzi, H. (2020). Land surface temperature variability across India: a remote sensing satellite perspective. *Theor. Appl. Climatol.*, *139*, 773-784,https://doi.org/10.1007/s00704-019-03010-8.

Raghuraman, S., Paynter, D., V., R., & Menzel, R. H. X. (2023). Greenhouse Gas Forcing and Climate Feedback Signatures Identified in Hyperspectral Infrared Satellite Observations. *Geophys. Res. Lett.*, *50*, https://doi.org/10.1029/2023GL103947.

Rodgers, C. (2000). Inverse Methods for Atmospheric Sounding. In (p. 256 pages). World Scientific, Singapore.

Roemer, F., Buehler, S., Brath, M., Kluft, L., & John, V. (2023). Direct observation of Earth's spectral long-wave feedback parameter. *Nature Geoscience*, *16*, 416-421,https://doi.org/10.1038/s41561-023-01175-6.

Sherwood, S., Ingram, W., Tshushuma, Y., Satoh, M., Roberts, M., Vidale, P., & O'Gorman, P. (2010). Relative humidity changes in a warmer climate. *J. Geophys. Res.*, *115*, https://doi.org/10.1029/2009JD012585.

Smith, N., & Barnet, C. (2019). Uncertainty Characterization and Propagation in the Community Long-Term Infrared Microwave Combined Atmospheric Product System (CLIMCAPS). *Remote Sensing*, *11*, 1227-1251,doi:10.3390/rs11101227.

Smith, N., & Barnet, C. (2020). CLIMCAPS observing capability for temperature, moisture, and trace gases from AIRS/AMSU and CrIS/ATMS. *AMT*, *13*, 4437–4459,https://doi.org/10.5194/amt-13-4437-2020.

Smith, N., & Barnet, C. (2023). Practical Implications of CLIMCAPS Cloud Clearing and Derived Quality Metrics. *Earth Space Sci*, *10*, e2023EA002913. https://doi.org/10.1029/2023EA002913.

Soden, B., & Held, I. (2006). An assessment of climate feedbacks in coupled ocean-atmosphere models. *J. Clim*, *19*, 3354-3360.

Strow, L., & DeSouza-Machado, S. (2020). Establishment of AIRS climate-level radiometric stability using radiance anomaly retrievals of minor gases and sea surface temperature. *Atmos. Meas. Tech.*, *13*, https://doi.org/10.5194/amt-13-4619-2020.

Strow, L., Hannon, S., DeSouza-Machado, S., Tobin, D., & Motteler, H. (2003). An overview of the AIRS radiative transfer model. *IEEE Transactions on Geosciences* *and Remote Sensing*, *41*, 303-313.

Strow, L., Hannon, S., Weiler, M., Overoye, K., Gaiser, S., & Aumann, H. (2003). Pre-launch spectral calibration of the Atmospheric InfraRed Sounder (AIRS). *IEEE Transactions on Geosciences and Remote Sensing*, *41*, 274-286.

Strow, L., Hepplewhite, C., Motteler, H., Buczkowski, S., & DeSouza-Machado, S. (2021). A Climate Hyperspectral Infrared Radiance Product (CHIRP) Combining the AIRS and CrIS Satellite Sounding Record. *Rem. Sens.*, *13*(3), https://doi.org/10.3390/rs13030418.

Susskind, J., Barnet, C., & Blaisdell, J. (2003). Retrieval of atmospheric and surface parameters from AIRS/AMSU/HSB data under cloudy conditions. *IEEE Transactions on Geosciences and Remote Sensing*, *41*, 390-409.

Susskind, J., Blaisdell, J., & Iredell, L. (2014). Improved methodology for surface and atmospheric soundings, error estimates, and quality control procedures: the Atmospheric Infrared Sounder science team version-6 retrieval algorithm. *J. App. Rem. Sens*, *8*(1), 084994 [10.1117/1.JRS.8.084994].

Tian, B., E., M., Roman, J., Thrastarson, H., & Fetzer, R., E. ad Monarrexz. (2020). *Airs version 7 level 3 product user guide* (Tech. Rep.). Jet Propulsion Laboratory, California Institute of Technology. Retrieved from `https://docserver.gesdisc.eosdis.nasa.gov/public/project/AIRS/V7_L3_User_Guide.pdf`

Whitburn, S., Clarisse, L., Bouilon, M., Safieddine, S., George, M., Dewitte, S., ... Clerbaux, C. (2021). Trends in spectrally resolved outgoing longwave radiation from 10 years of satellite measurements. *npj Clim Atmos Sci*, *4*, 48,https://doi.org/10.1038/s41612-021-00205-7.

Wu, W., Liu, X., Lei, X., L. amd Xiong, Yang, Q., Yue, Q., Zhou, D., & Larar, A. (2023). Single field-of-view sounder atmospheric product retrieval algorithm : establishing radiometric consistency for hyper-spectral sounder retrievals. *AMT*, *16*, 4807–4832, https://doi.org/10.5194/amt-16-4807-2023.

Zhou, D., Larar, A., Liu, X., Smith, W., Strow, L., Yang, P., ... Calbet, X. (2011). Global Land Surface Emissivity Retrieved From Satellite Ultraspectral IR Measurements. *IEEE Trans. Geosci. Remote Sens.*, *49*(4), 1277-1290.