https://doi.org/10.5194/egusphere-2024-3701
Preprint. Discussion started: 19 December 2024

Quantifying biases in TROPESS AIRS, CrIS, and joint AIRS+OMI tropospheric ozone products using ozonesondes

Elyse A. Pennington [1], Gregory B. Osterman[1], Vivienne H. Payne[1], Kazuyuki Miyazaki[1], Kevin W. Bowman[1], and Jessica L. Neu[1]

[1]Jet Propulsion Laboratory, California Institute of Technology

Correspondence: Elyse A. Pennington (elyse.a.pennington@jpl.nasa.gov)

Abstract. Quantifying changes in global and regional tropospheric ozone are critical for understanding global atmospheric chemistry and its impact on air quality and climate. Satellites now provide multi-decadal records of daily global ozone profiles, but previous studies have found large disagreements in satellite-based ozone trends, including in trends from different products based on the same spectral radiances. In light of these disagreements, it is critical to quantify to what degree the observed trend is attributable to measurement error for each product by comparing satellite-retrieved ozone to long-term measurements from ozonesondes. NASA's TRopospheric Ozone and its Precursors from Earth System Sounding (TROPESS) project provides satellite retrievals of ozone from a suite of instruments, including CrIS, AIRS, and multispectral combinations such as AIRS and OMI (joint AIRS+OMI) using a common algorithm. We compare these products to ozonesondes and find that the evolution of global tropospheric ozone satellite-sonde bias for TROPESS CrIS (0.21 ± 3.6% decade^{-1}, 2016-2021), AIRS (-0.41 ± 0.57% decade^{-1}, 2002-2022), and joint AIRS+OMI (1.1 ± 1.0% decade^{-1}, 2004-2022) is approximately one third the magnitude of trends in global tropospheric ozone reported by the Tropospheric Ozone Assessment Report Phase 1 (TOAR-I). We further quantify the bias in regional trends, which tend to be higher but with a smaller number of sondes, which can impact the satellite-sonde bias and trend. Our work represents an important basis for the utility of using satellite data to quantify changes in atmospheric composition in future studies.

1 Introduction

Tropospheric ozone trends show large regional variations and can depend strongly on the type of measurement analyzed, coverage, frequency, and vertical sensitivity, so there is no consensus about whether global tropospheric ozone is increasing or decreasing (Gaudel et al., 2018). The Tropospheric Ozone Assessment Report (TOAR, https://igacproject.org/activities/TOAR) aims to provide the most complete assessment of tropospheric ozone available to the community by compiling a comprehensive database of ozone measurements and analyzing data from multiple sources holistically. The satellite assessment in Phase 1 of TOAR (TOAR-I) demonstrated that an ensemble of satellite products agree on the present-day tropospheric ozone burden

https://doi.org/10.5194/egusphere-2024-3701
Preprint. Discussion started: 19 December 2024

(TOB) with a multi-satellite spread of only 37 Tg (approximately ±6%), but do not provide consensus on the long-term TOB trend (Gaudel et al., 2018). The six satellite products that they investigated spanned different time periods and displayed global TOB trends ranging from -2.15 Tg year^{-1} to 2.85 Tg year^{-1} (approximately -7.1% decade^{-1} to 9.5% decade^{-1}) (Table 1). While a thorough investigation of the origins of these disagreements was out of the scope of the report, speculated causes included differences in each instrument's vertical sensitivity to the atmospheric profile of ozone, spatial sampling across the globe, and measurement period.

Ongoing work in TOAR Phase 2 (TOAR-II) aims to investigate the causes of measurement-dependent differences in ozone trends and find consensus across techniques to determine the true TOB trend. There are a range of reported tropospheric ozone trends from work published in 2024 (Table 1), which agree that tropospheric ozone is increasing globally and in specific latitude bands, but to varying degrees and with varying levels of certainty. These recent works discuss two major reasons for discrepancies between satellite ozone products. The first is that the vertical sensitivity of each instrument impacts the amount of measured ozone in each level of the atmosphere (Gaudel et al., 2018). Some instruments and retrievals can distinguish influences from the upper versus lower troposphere (Pope et al., 2023, 2024; Froidevaux et al., 2024), which can lead to different assessments of how tropospheric ozone is changing if these changes are not uniform in altitude. Second, the satellite product quality can drift over time, producing an artificial trend caused by error in the instrument calibration or retrieval (Gaudel et al., 2018). Gaudel et al. (2024) demonstrated a method to correct the tropospheric ozone trend for the time-dependent bias, to produce a "true" tropospheric ozone trend. This was accomplished by quantifying the bias between satellite data and a reference method, for example ozonesonde data. The magnitude of bias was used to scale the satellite data to determine trends in tropospheric ozone with the retrieval bias (approximately) removed.

Ozonesondes provide a long-term record of detailed vertical ozone profiles and thus are an important tool for quantifying the vertical distribution of ozone and therefore have been useful in validating satellite retrievals of ozone. Ozonesondes are mounted on atmospheric balloons that are routinely launched from over 40 sites worldwide (WMO/GAW, 2024). These sites are distributed unevenly, with many sites located in North America and Europe, and fewer sites in the Southern Hemisphere. There are multiple types of ozonesondes, including electrochemical concentration cells (ECC; (Komhyr, 1969; Komhyr and Harris, 1971; Tarasick et al., 2021; WMO, 2021)) and Brewer-Mast (BM; (Brewer et al., 1997; Steinbrecht et al., 1998)) sondes. Most ozonesonde launch sites use ECC sondes, which typically have an uncertainty of 10-20% (WMO, 2021). The Harmonization and Evaluation of Ground-based Instruments for Free Tropospheric Ozone Measurements (HEGIFTOM) Working Group of TOAR-II was developed with the goal of improving the accuracy and precision of ozone measurements by removing inhomogeneities between ozonesondes caused by differences in equipment, operating procedures, or data processing. The Working Group utilized a harmonization technique to address this need and have reduced ozonesonde measurement uncertainty to 5% (5-10% in the tropics) (WMO, 2021). The state-of-the-art method produces sonde data with high accuracy, precision, and long-term reliability, making them a good tool for validating other measurement methods.

We introduce here three satellite tropospheric ozone products that were not part of the TOAR-1 assessment and use the HEGIFTOM ozonesonde measurements to assess their suitability for quantifying trends in TOB. The Tropospheric Ozone and its Precursors from Earth System Sounding (TROPESS) project (NASA, 2024) provides retrievals of ozone and other

https://doi.org/10.5194/egusphere-2024-3701
Preprint. Discussion started: 19 December 2024

Table 1. Tropospheric ozone burden (TOB) trends reported in the literature.

Product / Measurement	Trend (±Uncertainty)	Region	Date Range	Reference
Harmonized CCD	0.7 (±1.2) DU decade^{-1}	tropics	1995-2015	Heue et al. (2016)
Harmonized CCD	-0.2 (±0.6) DU decade^{-1}	tropics	1996-2015	Leventidou et al. (2018)
OMI/MLS	1.79 (±0.66) Tg yr^{-1}	60°S - 60°N	2005-2016	Gaudel et al. (2018)
IASI-FORLI	-2.15 (±1.03) Tg yr^{-1}	60°S - 60°N	2008-2016	Gaudel et al. (2018)
IASI-SOFRID	-1.34 (±0.92) Tg yr^{-1}	60°S - 60°N	2008-2015	Gaudel et al. (2018)
GOME/OMI	1.63 (±0.45) Tg yr^{-1}	60°S - 60°N	1996-2015	Gaudel et al. (2018)
OMI-RAL	2.85 (±1.16) Tg yr^{-1}	60°S - 60°N	2005-2015	Gaudel et al. (2018)
SCIAMACHY	1.50 (±1.39) Tg yr^{-1}	60°S - 60°N	2002-2012	Gaudel et al. (2018)
OMI/MLS	1-3 DU decade^{-1}	90°S - 90°N	2005-2016	Ziemke et al. (2019)
OMI/MLS	1.4 ppb decade^{-1}	90°S - 90°N	2005-2019	Fadnavis et al. (2024)
OMI/MLS	0.96 (±0.45) DU decade^{-1}	60°S - 60°N	2005-2019	Elshorbany et al. (2024)
OMI/MLS	1.06 (±0.40) DU decade^{-1}	30°S - 30°N	2005-2019	Elshorbany et al. (2024)
OMI/MLS	0.78 (±1.16) DU decade^{-1}	30°N - 60°N	2005-2019	Elshorbany et al. (2024)
OMI/MLS	0.95 (±0.75) DU decade^{-1}	30°S - 60°S	2005-2019	Elshorbany et al. (2024)
OMI/MLS	1.6 (±1.1) Tg decade^{-1}	0°- 20°N	2004-2019	Gaudel et al. (2024)
OMI	2.4 (±1.1) Tg decade^{-1}	0°- 20°N	2004-2019	Gaudel et al. (2024)
OMI/MLS	0.9 (±2.2) Tg decade^{-1}	0°- 20°S	2004-2019	Gaudel et al. (2024)
OMI	1.9 (±2.4) Tg decade^{-1}	0°- 20°S	2004-2019	Gaudel et al. (2024)

trace gases from a suite of satellite instruments, including the Cross-track Infrared Sounder (CrIS), Atmospheric Infrared Sounder (AIRS), and a joint retrieval product using AIRS and the Ozone Monitoring Instrument (OMI) (joint AIRS+OMI). These satellite products provide long-term records of ozone using a consistent retrieval algorithm that produces ozone profiles that use the same *a priori* profiles and are calculated on the same vertical grid and with the same method of uncertainty estimation, making them more readily comparable. This study aims to validate the accuracy of TROPESS satellite retrievals of tropospheric ozone and their stability with time against ozonesonde measurements. Section 2 introduces the satellite and ozonesonde datasets and describes the analysis tools used to compare them. Section 3 presents comparisons between satellite and ozonesonde profiles and columns, long-term trends in the satellite-sonde bias, and temporal and geographic variations in these quantities.

https://doi.org/10.5194/egusphere-2024-3701
Preprint. Discussion started: 19 December 2024

2 Data and methods

2.1 TROPESS ozone retrievals

All TROPESS Level 2 data is produced using the MUSES — MUlti-SpEctra, MUlti-SpEcies, MUlti-SEnsors — algorithm (Fu et al., 2013, 2016, 2018) following the optimal estimation methods employed for the EOS Tropospheric Emission Spectrometer (TES) (Beer, 2006; Bowman et al., 2006). This approach estimates vertical profiles, uncertainty estimates, and observation operators, which are critical for chemical data assimilation and inverse modeling (Jones et al., 2003; Parrington et al., 2009; Miyazaki et al., 2017, 2020b). Three TROPESS products are considered in this study, CrIS, AIRS, and the joint AIRS+OMI retrieval, and have been used in previous studies to, for example, monitor changes in global tropospheric ozone (Miyazaki et al., 2021) and understand the processes controlling air pollution (Miyazaki et al., 2019).

CrIS, AIRS, and OMI data have been used to generate Level 2 products by other teams and retrieval algorithms. For example, the Community Long-term Infrared Microwave Combined Atmospheric Product System (CLIMPCAPS), NOAA Unique Combined Atmospheric Processing System (NUCAPS), and Near Real Time (NRT) Standard Physical Retrieval (AIRS2RET_NRT_7.0) produce Level 2 CrIS and AIRS products that are retrieved from Level 2 cloud cleared radiances on a 45 km field of regard (Smith and Barnet, 2020; Barnet et al., 2021; AIRS, 2019). The single field-of-view (SFOV) sounder atmospheric product (SiFSAP) retrieval algorithm produces Level 2 CrIS, AIRS, and IASI products from Level 1B radiances (Wu et al., 2023). The TROPESS project generates Level 2 products using Level 1B radiances processed by the MUSES algorithm. The three TROPESS products are described further in Sect. 2.1.1-2.1.3.

The global, monthly average tropospheric ozone columns are computed from the ozone profiles retrieved by MUSES and are shown in Fig. 1 for the three products. The average tropospheric column — integrated over the ozone profile from the surface to the thermal tropopause (see Sect. 2.3) — falls between 30 and 40 Dobson Units (DU), with the ozone seasonal cycle clearly visible. Each satellite product provides data for different time periods and with different average magnitudes, which will be discussed in Sect. 3.

2.1.1 CrIS

The Cross-track Infrared Sounder (CrIS) is on the Suomi National Polar-orbiting Partnership (SNPP) satellite, which was launched on October 28, 2011 and has a 13:30 local time ascending node (Han et al., 2013). CrIS instruments are also flying on the Joint Polar Satellite Series platforms (JPSS-1 and JPSS-2). However, for this study we focus on SNPP-CrIS data. CrIS has high spectral resolution, high signal-to-noise ratio, and a low calibration uncertainty (Han et al., 2013; Strow et al., 2013; Tobin et al., 2013; Wang et al., 2013). CrIS measures infrared (IR) radiances in 3 bands: 650-1090 cm^{-1} (long-wave IR, LWIR), 1210-1750 cm^{-1} (mid-wave IR, MWIR), and 2155-2550 cm^{-1} (short-wave IR, SWIR) (Han et al., 2013). TROPESS ozone retrievals from CrIS specifically use 4 windows in the LWIR and 6 windows in the MWIR (Malina et al., 2024). TROPESS CrIS Level 2 products are retrieved from the L1B radiances on individual 15 km fields of view. The CrIS data record begins in 2012 but the radiance data that was initially retrieved only has nominal spectral resolution (NSR): 0.625 cm^{-1} in the LWIR band, 1.25 cm^{-1} in the MWIR band, and 2.5 cm^{-1} in the SWIR band (Han, 2015). In November 2015, the data record grew

https://doi.org/10.5194/egusphere-2024-3701
Preprint. Discussion started: 19 December 2024

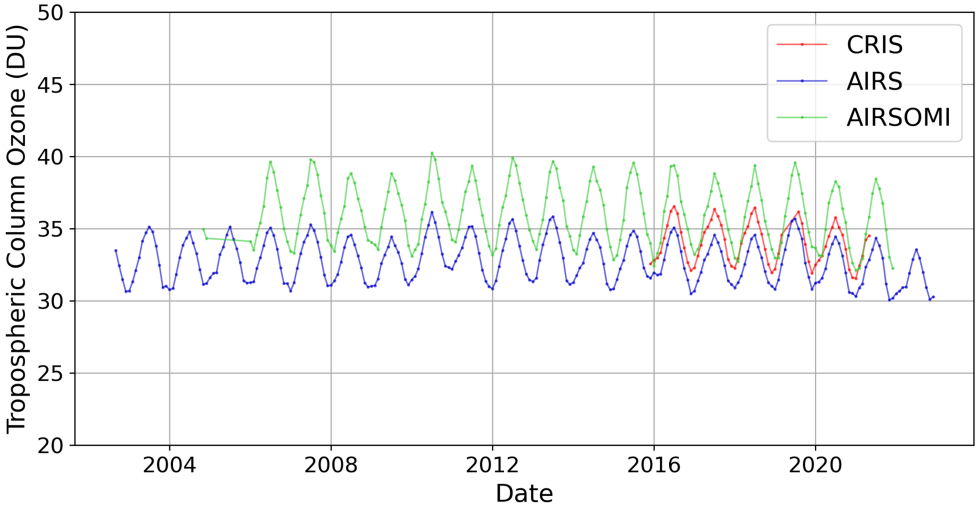

Figure 1. Global (67°S-85°N) monthly average tropospheric ozone columns for the TROPESS CrIS (red), AIRS (blue), and joint AIRS+OMI (green) products.

to include the full spectral resolution (FSR) of 0.625 cm^{-1} in all three bands. The TROPESS project uses the NASA L1B FSR radiance data (Revercomb and Strow, 2018) so our record begins in November 2015 (Han, 2015). In late March 2019, there was an anomaly in the MWIR band, resulting in a gap in data from April through July 2019. The instrument was restored to full functionality, and there is good consistency between the early and later 2019 data (Iturbide-Sanchez et al., 2022). In July 2021, there was a failure in the LWIR bands that required the instrument to change to a different set of electronics (NOAA, 2021), and the data acquired after this date was no longer included in the long-term record by NASA. Therefore, the ozone record used in this study runs from December 2015 through May 2021. Given the large variability in tropospheric ozone and the small magnitude of its trends, the TROPESS CrIS record is too short to detect ozone trends with good statistical significance. Nonetheless, we consider the time dependence of the CrIS-sonde bias using the same analysis process as with the other two TROPESS products because we are focused here on whether there are any changes in the retrieval quality with time as well as on the consistency of the TROPESS products. In particular, TROPESS CrIS and AIRS both utilize IR radiances, and their comparison is of particular interest. We discuss the impact of the relatively short CrIS time range with respect to bias detection in Sect. 3.2. Our analysis is a necessary step if the TROPESS CrIS product is to be used as part of a longer-term merged tropospheric ozone record in the future.

2.1.2 AIRS

The Atmospheric Infrared Sounder (AIRS) is a thermal infrared (TIR) grating spectrometer onboard the Aqua satellite, launched on May 4, 2002 (Aumann et al., 2003; Susskind et al., 2003, 2014). Aqua is part of the A-train, a constellation of satellites that orbit together (Stephens et al., 2002). AIRS has a 13 km footprint at nadir and measures in the 650-2675 cm^{-1}

https://doi.org/10.5194/egusphere-2024-3701
Preprint. Discussion started: 19 December 2024

wavelength range (Aumann et al., 2003; Susskind et al., 2003), with ozone retrieved from channels around 1042 cm^{-1} (Wei et al., 2010). Similarly to CrIS, TROPESS retrieves ozone from AIRS from L1B radiances (AIRS, 2007) using 10 distinct wavelength windows in the IR (Fu et al., 2018). TROPESS AIRS Level 2 products are retrieved from the L1B radiances on individual 15km fields of view.

2.1.3 Joint AIRS+OMI retrieval

The Ozone Monitoring Instrument (OMI), unlike CrIS and AIRS, measures ozone-absorbing radiances in the ultraviolet (UV) range (Levelt et al., 2006, 2018; Liu et al., 2010b, a; Huang et al., 2017). OMI flies on the Aura satellite, a member of the A-train, so it collects data at nearly the same local time as AIRS. Retrievals of ozone from OMI use wavelengths between 270 and 365 nm with a ground pixel size of 12 x 24 km (Levelt et al., 2006). OMI experienced a "row anomaly" beginning in 2007. In 2009, the anomaly spread and many pixels became unusable; since 2010, ~40-50% of OMI pixels must be discarded. In addition to the lower data volume, there have been small changes to OMI radiances (1-2%) and irradiances (3-8%), but overall the good pixels remain stable (Schenkeveld et al., 2017).

Fu et al. (2018) combined AIRS and OMI radiances in a joint retrieval to produce the joint AIRS+OMI product. The multispectral retrieval, implemented in the MUSES algorithm, uses information in the AIRS TIR bands and the OMI UV bands to produce vertical ozone profiles. The joint AIRS+OMI product has greater vertical sensitivity than the AIRS or OMI retrievals alone, with degrees of freedom of signal (DOFS) in the troposphere ranging from 0.2 to 1.6, typically falling above 1 (Fu et al., 2018). Because the DOFS exceeds 1, it is possible to resolve ozone subcolumns in the troposphere, i.e., upper and lower tropospheric ozone.

2.2 Ozonesonde data

Ozonesondes measure the partial pressure of ozone and the total atmospheric pressure, while a radiosonde measures atmospheric temperature. These quantities can be used to calculate the volume mixing ratio (VMR) of ozone throughout the profile until the balloon pops, typically around 80 hPa. Ozonesonde data is provided by the HEGIFTOM Working Group of TOAR-II (HEGIFTOM, 2024; Van Malderen, 2023). The sonde launch sites providing data for this study are shown in Fig. 2a.

2.2.1 Coincidence criteria

Sonde and satellite profiles are compared if the sonde is launched within 300 km and 9 hours of the satellite measurement. These criteria may result in multiple satellite profiles being matched with a single sonde profile, so repeat satellite-sonde matches are removed by only considering the match with the lowest distance or time offset between the satellite overpass and the sonde launch. These spatial and temporal thresholds were chosen to compare measurements from similar air masses while being large enough to capture sufficient data for a statistically meaningful comparison, and to maintain consistency with previous studies (Nassar et al., 2008; Verstraeten et al., 2013). However, the choice of coincidence criteria can impact the satellite-sonde bias. Nassar et al. (2008) showed that wide spatial and temporal coincidence criteria can lead to comparisons between satellite and

https://doi.org/10.5194/egusphere-2024-3701
Preprint. Discussion started: 19 December 2024

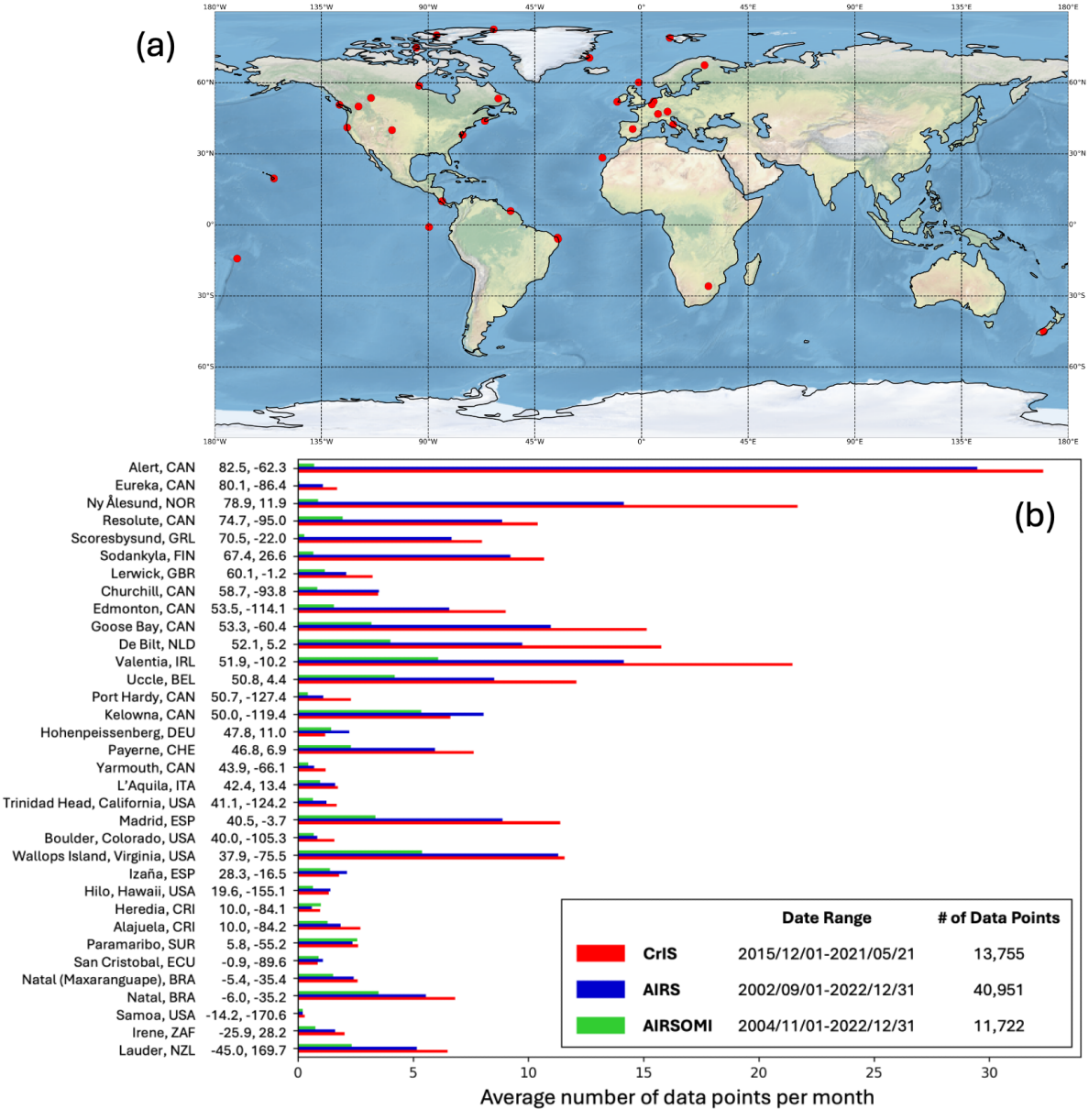

Figure 2. (a) Ozonesonde launch sites providing harmonized data in this study. (b) Monthly average number of matched satellite-sonde data points at each sonde launch site for each satellite product. Sites are ordered by decreasing latitude. (b), inset: The date range of each satellite product and the total number of matched satellite-sonde data points over the date range at all sites.

sonde profiles that measured air columns with high atmospheric and ozone variability. When they tightened the spatial and temporal coincidence criteria, the value of the bias did not change in a statistically meaningful way, but the standard deviation increased. This implies that comparing profiles with different atmospheric profiles adds random error but does not introduce

https://doi.org/10.5194/egusphere-2024-3701
Preprint. Discussion started: 19 December 2024

positive or negative bias (Nassar et al., 2008). So, we maintain these coincidence criteria — 300 km and 9 hours — in our study.

2.2.2 Application of the satellite operator

To directly compare satellite and sonde profiles, the satellite retrieval operator must be applied to the sonde data to account for satellite sensitivity and vertical resolution (Jones et al., 2003; Worden et al., 2007; Verstraeten et al., 2013; Malina et al., 2024). This operation is performed on the sonde data to estimate the ozone profile that the satellite would have measured had it observed the same air mass as the sonde, with no systematic or random errors other than those attributable to the smoothing of the profile due to the vertical sensitivity of the satellite. The satellite operator includes the averaging kernel and *a priori* constraint vector together; examples of these quantities are shown in Fig. 3. The following steps describe the application of the satellite operator to a sonde profile.

1. Sonde measurements are provided on fine vertical grids with variable maximum altitude. All TROPESS profiles are provided on 67 fixed levels reaching a minimum pressure of approximately 1 hPa. For a matched pair of a satellite profile and sonde profile, if the minimum pressure (i.e., maximum altitude) of the sonde profile is greater than the minimum pressure of the satellite profile, then the ozone concentrations from the satellite prior are appended to the sonde profile concentrations, scaling the prior so that the concatenated concentration profile is continuous. This step ensures that the satellite and sonde profiles reach the same minimum pressure.

2. The sonde profile is interpolated to the same 67 pressure levels as the satellite profile.

3. The sonde operator is applied according to Equation 1. $\boldsymbol{x}_{sonde}$ is the new sonde profile, which can be directly compared to the satellite profile, $\boldsymbol{x}_a$ is the *a priori* profile, $\boldsymbol{A}$ is the averaging kernel, and $\boldsymbol{x}_{sonde,measured}$ is the sonde profile produced following steps 1 and 2. Example profiles and averaging kernel from a single CrIS sounding are presented in Fig. 3, and representative averaging kernels for each instrument are given in Fig. S4.

$$\boldsymbol{x}_{sonde} = \boldsymbol{x}_a + \boldsymbol{A}(\boldsymbol{x}_{sonde,measured} - \boldsymbol{x}_a) \tag{1}$$

After the satellite operator is applied to the sonde profile, the profiles must be quality controlled. Poor quality in the resampled sonde profiles may result from either poor quality in the original sonde profiles ($\boldsymbol{x}_{sonde,measured}$) or poor behavior when the satellite operator is applied to $\boldsymbol{x}_{sonde,measured}$. The HEGIFTOM group did not quality control (QC) the sonde data, so we tested multiple methods of QC, explained in Sect. S1. To evaluate each QC method, we considered the impact on the mean, median, standard deviation, and trend of the tropospheric ozone column satellite-sonde bias, as well as the percentage of profiles removed. See Sect. S1 for a detailed description of the methods that were tested and the metrics used to evaluate the QC methods.

Our selected quality control method considers the stratosphere and troposphere separately. In the stratosphere, profiles are compared to climatology from the Aura Microwave Limb Sounder (MLS), an instrument launched in 2004 onboard the Aura satellite that provides stratospheric profiles of multiple atmospheric gases (Livesey et al., 2022; Waters et al., 2006; Werner

https://doi.org/10.5194/egusphere-2024-3701
Preprint. Discussion started: 19 December 2024

et al., 2023) and has been validated using ozonesonde and lidar measurements (Jiang et al., 2007). The Level 3 MLS dataset includes the mean ozone profile and standard deviation at each level, binned into latitude bands spaced every 10° (Fig. S5). We assume that any ozone mixing ratio that falls outside of the mean MLS profile ±5 standard deviations in a specific latitude band is physically impossible, so if a sonde or satellite VMR falls outside of these limits at any level, that profile is removed from the analysis. MLS only measures ozone in the upper troposphere and above, so the ozonesonde concentrations within the troposphere are compared to the distribution of the TROPESS satellite ozone profiles. For each satellite-sonde set (e.g., AIRS + sondes), each sonde profile is compared to the mean ±5 standard deviations of the distribution of satellite profiles. If any concentrations in the sonde profile fall outside of this range, the profile is removed from the analysis. This process is performed in the troposphere for each set of matched satellite-sonde profiles. Using this method of QC, the percentage of profiles removed from CrIS, AIRS, and joint AIRS+OMI comparisons are 4.5%, 3.5%, and 5.7%, respectively.

The number of matched, filtered data points for each satellite product at each sonde launch site is given in Fig. 2b. The differences in the number of matched data points between satellite products is due to the differences in sampling, satellite data throughput, and time period of data collection (see Sect. 2.1.1-2.1.3).

2.3 Column calculations

The MUSES algorithm method is used to calculate total ozone columns and ozone subcolumns. The total ozone column is calculated by integrating the ozone from the surface to the top of the atmosphere — the highest-altitude point in the TROPESS products, around 0.1 hPa. Calculating a subcolumn requires the definition of the maximum and minimum pressure levels for the column, as well as temperature and water vapor profiles to account for factors such as the Bernoulli equation and deviation from the ideal gas law. Temperature and water vapor profiles are provided by the TROPESS project (NASA, 2024).

Tropospheric columns are integrated from the surface to the thermal tropopause, as defined by the World Meteorological Organization (WMO) (WMO, 1957). The thermal tropopause is obtained from version 5 of the NASA Global Modeling and Assimilation Office (GMAO) Goddard Earth Observing System (GEOS-5) model Forward Processing for Instrument Teams (FP-IT) (Molod et al., 2012; GMAO, 2024; Lucchesi, 2015) and is specific to the date, time, and location of each data point. TROPESS defines the lower tropospheric column to extend from the surface to 500 hPa, and the upper tropospheric column to extend from 500 hPa to the tropopause.

2.4 Correlation and trend analysis

The percent bias is calculated for every profile, column, and subcolumn according to Equation 2,

$$Percent\ Bias = \frac{C_{satellite} - C_{sonde}}{C_{sonde}} * 100\% \quad (2)$$

The calculation of trends follows guidance presented by the TOAR-II statistics focus working group (Chang et al., 2023). In brief, we use quantile regression (QR) to report trends at the 50th percentile. In comparison to other trend detection methods, QR is preferred due to its robustness to small sample sizes and outliers and its ability to account for non-normal error distri-

https://doi.org/10.5194/egusphere-2024-3701
Preprint. Discussion started: 19 December 2024

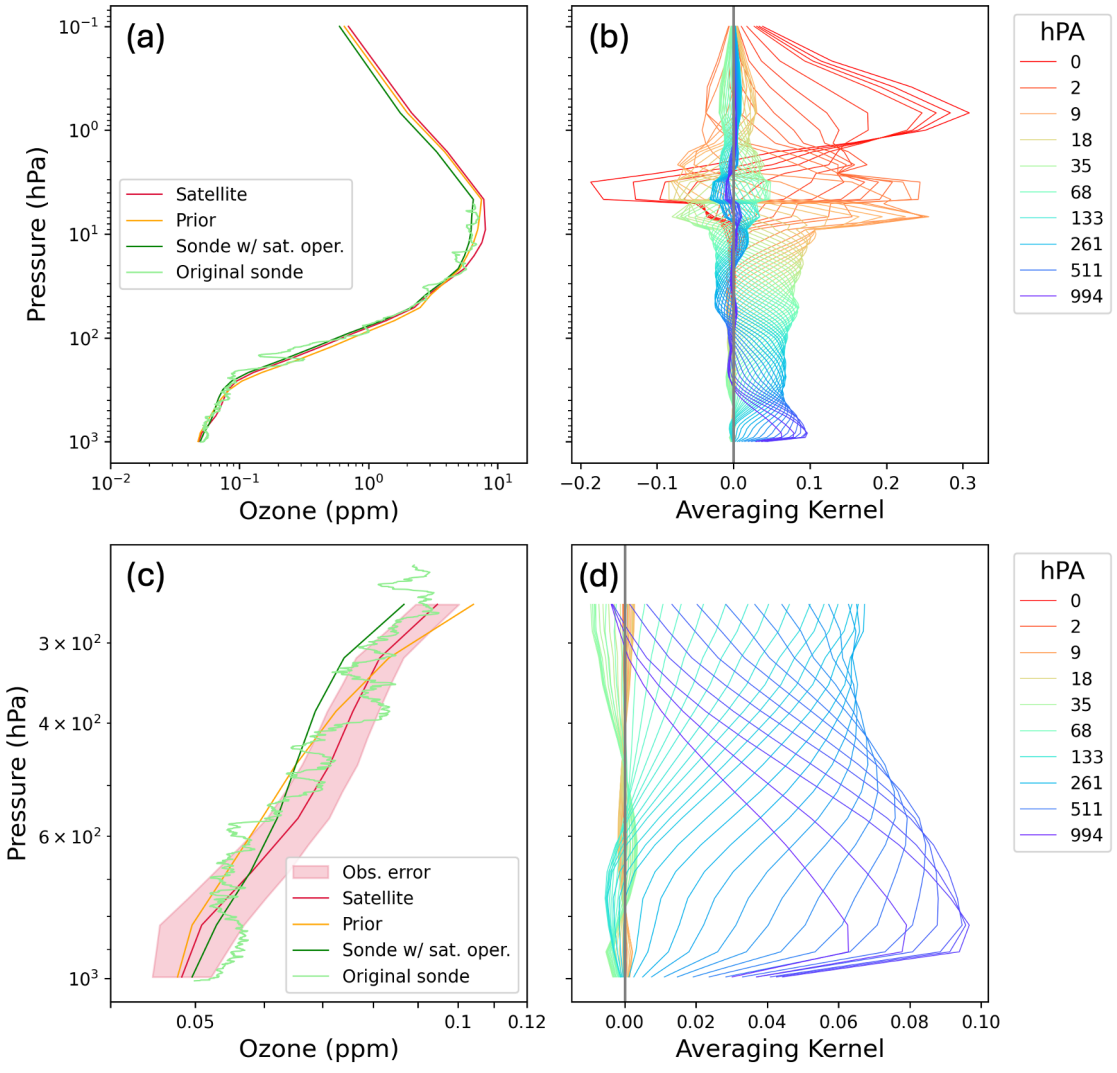

Figure 3. Example ozone profiles and averaging kernel from a comparison of a CrIS retrieval to an ozonesonde. (a) Ozone profile retrieved from the satellite (red), prior profile used in the CrIS retrieval (orange), profile measured by the ozonesonde (light green), and profile produced when the satellite operator is applied to the ozonesonde measurement (dark green). (b) Averaging kernel corresponding to the same retrieval in A. (c,d) Same as (a,b), zoomed into the troposphere. The red shading in C indicates the observation error of the satellite profile.

butions and autocorrelation. The TOAR-II guidance also provides a method for moving block bootstrapping to calculate the uncertainty of trends and assign a p-value to communicate the likelihood that a trend exists. Moving block bootstrapping calculates the standard error of the trend over multiple subsamples ("blocks") along a time series, and is used to accurately quantify the trend uncertainty while taking autocorrelation and heteroskedasticity into account (Chang et al., 2023). Satellite-sonde correlation coefficients (r^2) are calculated using simple linear regression.

https://doi.org/10.5194/egusphere-2024-3701
Preprint. Discussion started: 19 December 2024

To explore the geographic variability of ozone levels and data quality, results are grouped by latitude bands that are spaced every 30°.

3 Results

3.1 Satellite-sonde comparisons

The satellite and sonde profiles matched for each satellite product (Fig. S6) were used to derive profiles of percent difference for the three TROPESS products globally and within each latitude band (Fig. 4). The global mean percent bias profiles all fall within 20% bias, with the vertical distribution of bias varying by product. Each satellite instrument has a different vertical sensitivity, as indicated by the averaging kernels (Fig. 3 and S4), which results in differences in the ozone vertical profile. However, because each satellite operator has been applied to each sonde profile, the biases primarily reflect differences in the air mass sampled within the coincidence criteria along with systematic and random errors in the satellite retrievals. The magnitude and vertical distribution of bias varies across latitude bands, but there is more variability across instrument than across latitude for each instrument. Figure 4 illustrates the high volume of data in the Northern mid and high latitudes compared to the tropics and Southern Hemisphere, due to the lack of ozonesonde data in the Southern Hemisphere.

The TOB is represented by tropospheric column ozone (Fig. 1), so we calculate the percent bias between matched satellite- and sonde-measured columns. A time series of monthly averaged percent bias between matched satellite and sonde tropospheric ozone columns between 30 and 60°N is shown in Fig. 5a. The 50th percentile (i.e., median) of all of the points is given in the legend for each satellite product, and is displayed in the left-hand bar chart (Fig. 5b) for all latitude bands. All three satellite products tend to overestimate tropospheric ozone columns in all latitude bands, with AIRS having the lowest positive bias. Joint AIRS+OMI tends to have the largest bias, but the CrIS error can be comparable to joint AIRS+OMI. The median percent bias at each sonde site is given in Fig. S7. The median bias is positive at most sites, but there are some negative median bias values. There can be large differences in the bias of the three products at some sites, and even sites within the same latitude band do not necessarily show consistent biases. The same analysis as shown in Fig. 5 but using absolute bias (in Dobson Units) is shown in Fig. S8.

To better understand the differences in satellite bias between products, the 50th percentile percent bias for each latitude band in both the lower troposphere (LT) and upper troposphere (UT) is given in Fig. 6. The largest overpredictions are in the UT, where satellite sensitivity tends to be larger (Fig. S4). The distribution of bias in the UT across latitude bands is consistent with the overall tropospheric column bias. In the LT, AIRS' and CrIS' error fluctuates around 0. AIRS and CrIS have lower sensitivities in the LT, so there is little true information in this altitude range and the lack of bias primarily reflects the use of the same prior profile in both the satellite measurements and the sonde measurements with the satellite operator applied. Joint AIRS+OMI has greater sensitivity in the LT than the other two products, and it also tends to have the largest magnitude bias; it is negative at high northern latitudes, but positive in all other regions.

The TROPESS products' tropospheric ozone biases are similar to the bias between other satellite products and ozonesondes. The mean tropospheric column ozone bias between the Tropospheric Emission Spectrometer (TES) and ozonesondes

https://doi.org/10.5194/egusphere-2024-3701
Preprint. Discussion started: 19 December 2024

was shown to range from 7 to 15% (Verstraeten et al., 2013) or approximately 2.9 DU (Osterman et al., 2008), with the UT displaying a wider range of bias than the LT (Nassar et al., 2008; Verstraeten et al., 2013). Our results show a decreased tropospheric column bias compared to TES (Fig. 5b), with a consistent distribution of error between the UT and LT compared to TES. Whereas the three satellite products studied in this work display positive mean biases in most locations, other satellites display negative bias in some latitude bands. Boynard et al. (2018) showed that the IASI (Infrared Atmospheric Sounding Interferometer)/Metop-A (IASI-A) and IASI/Metop-B (IASI-B) tropospheric ozone products had mean biases compared to sondes ranging 4-5% in high latitudes, -4 to -5% in midlatitudes, and -16 to -19% in the tropics. Miles et al. (2015) showed a mean bias of approximately 10% (2 DU) in the Northern Hemisphere and -15 to -20% (-1 to -3 DU) in the Southern Hemisphere for the Global Ozone Monitoring Experiment 2 (GOME-2) RAL (Rutherford Appleton Laboratory) ozone retrieval. The OMI/MLS product presented by Ziemke et al. (2019) reported a mean bias of -2 DU for the tropospheric column ozone compared to global ozonesondes between 60°N and 60°S. While the sign of the mean biases can vary between satellite products, the absolute values are of approximately the same order of magnitude, suggesting that the TROPESS products are of a similar quality as existing satellite products. Gaudel et al. (2018) calculated a mean TOB of 301 Tg with a range of 281-318 Tg across five satellite products. The 37 Tg range is equivalent to approximately ±6% of the average TOB. The differences in satellite-sonde bias across satellite products, ranging from -3 to 3 DU as shown in our results and the previous studies summarized here, likely accounts for at least some of the difference in measured TOB presented in Gaudel et al. (2018).

3.2 Trends in satellite-sonde bias

The time series of tropospheric ozone column biases (such as that shown in Fig. 5a) is used to calculate the trend in satellite-sonde bias and its uncertainties. The trends for each product are shown in the legend of Fig. 5a and for all latitude bands in Fig. 5c. Figure S8 displays the trends in concentration units (DU decade^{-1}) and Fig. S7 displays the trends for each sonde launch site. The trend in satellite-sonde bias is within ±2% decade^{-1} for all products globally. Gaudel et al. (2018) reported TOB trends ranging from -2.15 Tg year^{-1} to 2.85 Tg year^{-1}, which equates to approximately -7.1% decade^{-1} to 9.5% decade^{-1} with the reported mean global ozone burden of 301 Tg. Therefore, the TROPESS products exhibit global trends in bias about 3 times smaller than currently reported tropospheric ozone trends.

When reported in each latitude band, the magnitude of the bias trend tends to be larger and the sign can vary. In most regions, the standard error on the trend (represented by the black lines in Fig. 5c) is larger than the trend itself and the p-values are large (Table S4). This suggests that the trends have very low certainty and cannot be distinguished from a trend of zero, and thus should not influence TOB trend detection. This is not true in the tropics, however, where there are relatively larger trends and smaller errors. This feature is investigated further in Sect. 3.3. A previous study comparing IASI-A and IASI-B to ozonesondes found an average drift of -8.6 (±3.4) % decade^{-1} (-2.81 (±1.26) DU decade^{-1}) in the Northern Hemisphere. The TROPESS products all have drift values less than that value, excluding CrIS in the tropics. Because most of the sonde data is collected in Northern mid-latitudes, the global average satellite-sonde bias is most strongly influenced by the data in those latitude bands.

To further investigate the spatial distribution of errors in satellite retrievals, we quantify the correlation between the satellite-sonde biases at different sonde launch sites. If the monthly time series of percent bias at different sites are uncorrelated, then

https://doi.org/10.5194/egusphere-2024-3701
Preprint. Discussion started: 19 December 2024

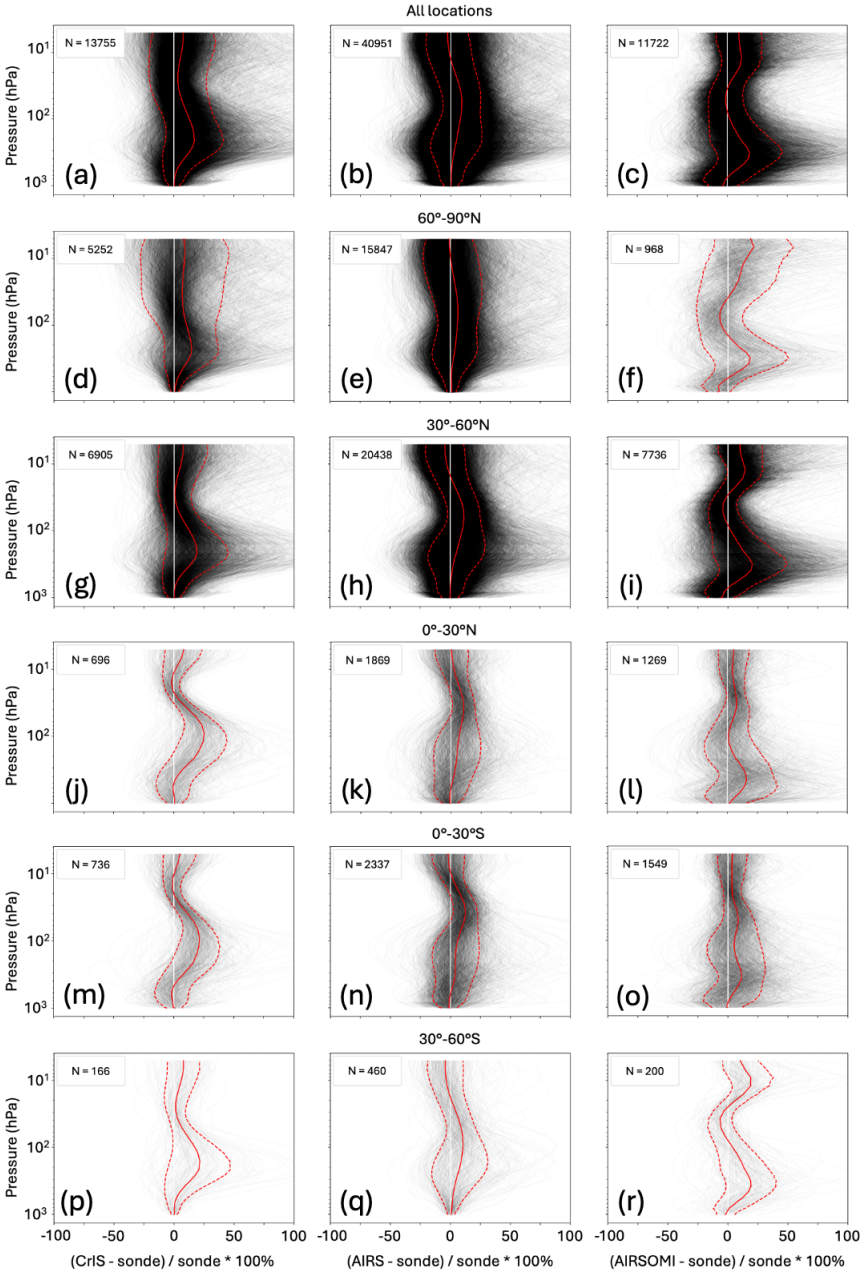

Figure 4. Percent bias between each matched satellite and sonde profile for CrIS (left), AIRS (center), and joint AIRS+OMI (right). The top row includes data at all sites, and subsequent rows contain data in each 30°latitude band. The solid red lines represent the mean difference profile and the dashed red lines represent one standard deviation away from the mean. The number of profiles in each latitude band is indicated by the density of the individual black lines. The solid white lines are at 0%.

https://doi.org/10.5194/egusphere-2024-3701
Preprint. Discussion started: 19 December 2024

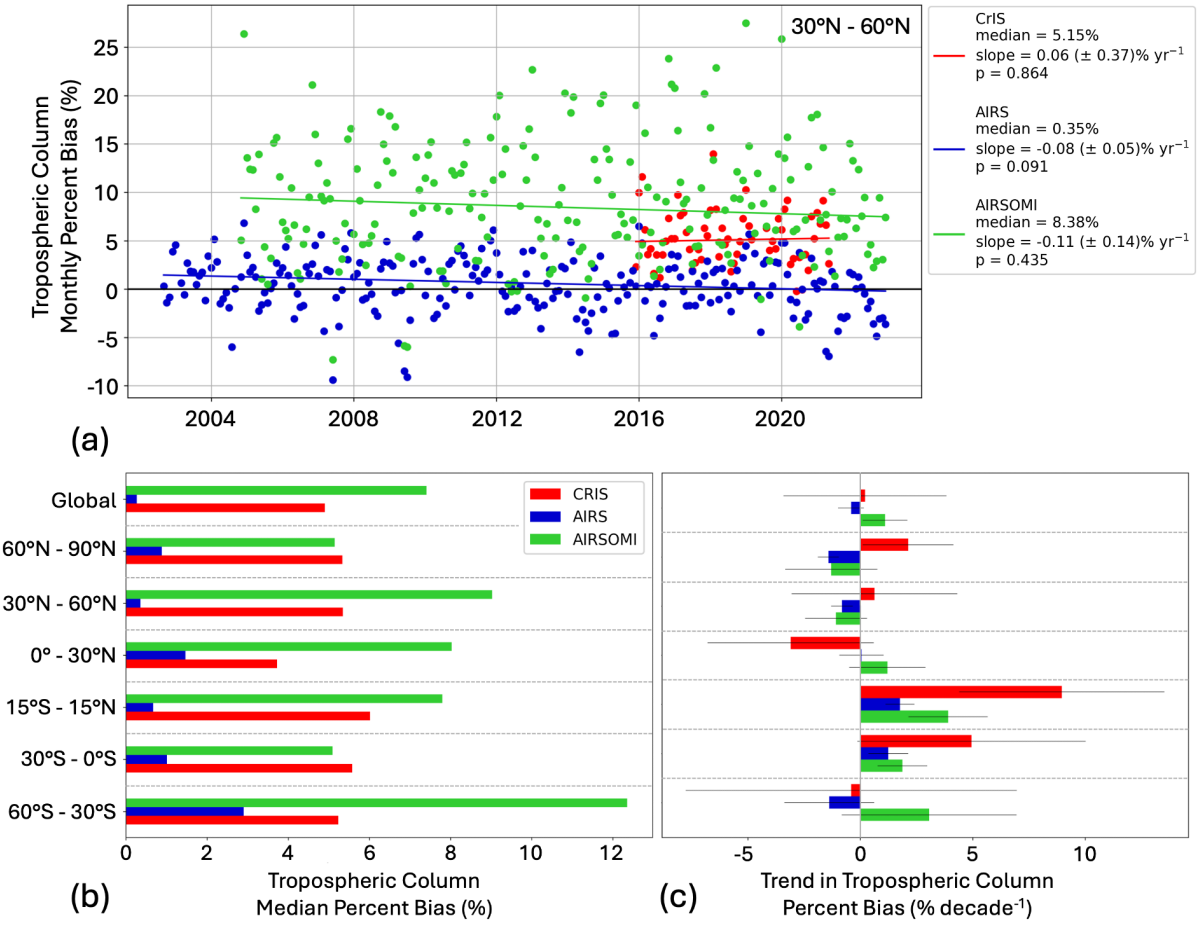

Figure 5. Top: Monthly-averaged tropospheric column percent bias for all sites between 30°N and 60°N. The slope, standard error of the slope, and p-value of the slope are reported for the 50th percentile. The medians and slopes reported in the bottom row are derived from time series such as this. Bottom left: The median tropospheric column percent bias in each latitude band. Bottom right: The trend in tropospheric column percent bias in each latitude band. There are no sonde sites between 60°S and 90°S.

the satellite products are providing geographically independent data. As an example, the bias time series at one site (e.g., Uccle, Belgium) was compared to the bias time series at a second site (e.g., Valentia, Ireland), and the r^2 between the time series data was calculated. This analysis was performed between all pairs of sites and the r^2 values are given in Fig. S9. High r^2 values indicate a relationship between the bias in different locations, while a low r^2 value indicates no relationship. The r^2 values are consistently low between sites, even sites that are in similar regions such as mainland Europe (De Bilt, Valentia, Uccle, Hohenpeissenberg, Payerne, L'Aquila, and Madrid), Costa Rica (Alajuela and Heredia), or Brazil (Natal and Maxaranguape). This suggests that satellite-sonde bias is random and not systematic by location. While the sonde sites do not represent the

https://doi.org/10.5194/egusphere-2024-3701
Preprint. Discussion started: 19 December 2024

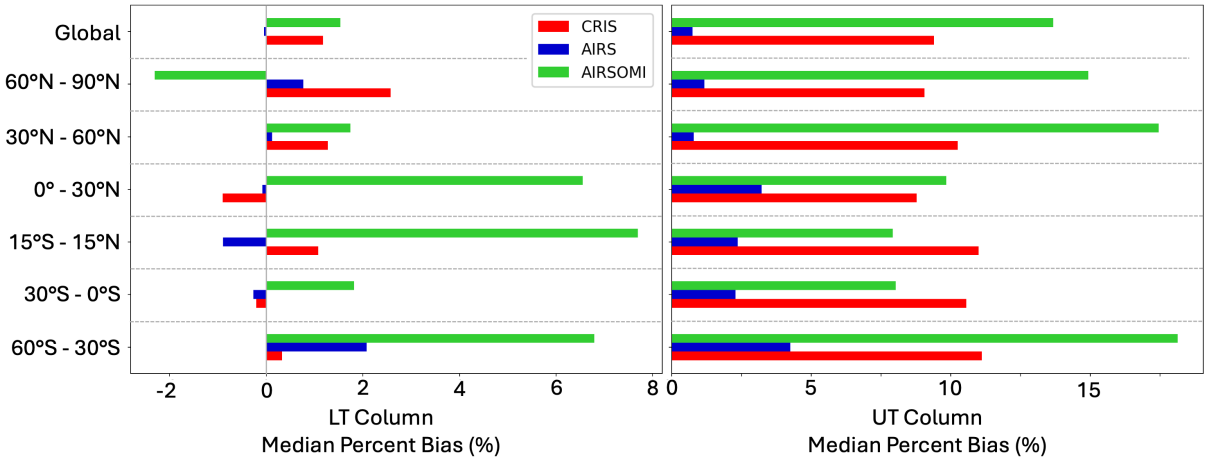

Figure 6. The median LT (left) and UT (right) column percent bias averaged over sites in each latitude band. There are no sonde sites between 60°S and 90°S.

entire globe and are not evenly distributed, this preliminary analysis suggests that the satellite performance is not strongly location-dependent.

The satellite products cover different time periods, with CrIS providing data for the shortest time period. To compare the products during only their overlapping time periods, the AIRS and joint AIRS+OMI data were cropped to the same time period as CrIS and the median percent bias and the trend in the percent bias were calculated in the same manner as Fig. 5 (see Fig. S10). The change in the median percent bias between the two time periods was minor at almost all sites, but the magnitude of the trend in the bias increased at many sites. At the same time, the standard errors of the trends tend to be larger than the trends themselves and the p-values are large (Fig. S10), meaning that the trends in bias during this time period are uncertain and may not be differentiable from 0. The large errors highlight the difficulty of trying to determine a trend over a short time period, particularly for a trace gas with large interannual variability like tropospheric ozone. While one may be tempted to infer that these 2015-2021 bias trends indicate the current performance of the AIRS and joint AIRS+OMI satellite products, the uncertainty of the trends suggests rather that 5 years is not a sufficient period for trend detection. The results also imply that the relatively large trends in bias seen in CrIS as compared to the other two instruments are primarily driven by the short length of the record rather than by any instability or deficiency in the CrIS product.

3.3 Seasonal dependence of biases

The seasonal dependence of tropospheric, UT, and LT column satellite-sonde comparisons is shown in Fig. 7. Seasonal dependencies arise in CrIS and AIRS because they are both TIR instruments and thus their sensitivity to ozone depends on the thermal structure of the atmosphere, particularly the thermal contrast between the surface and the lowest layer of the atmosphere. The sensitivity of OMI likewise varies with season, as it depends strongly on solar insolation. As expected, AIRS and

https://doi.org/10.5194/egusphere-2024-3701
Preprint. Discussion started: 19 December 2024

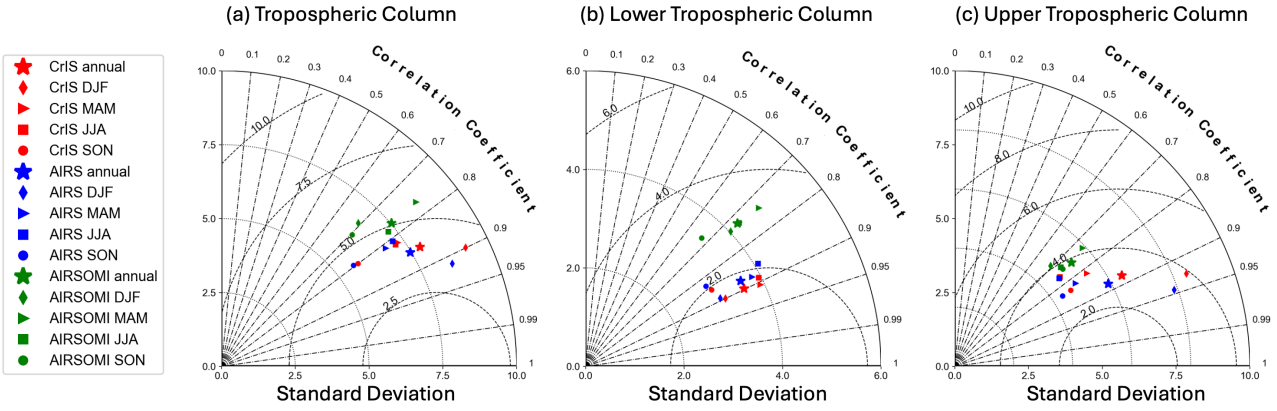

Figure 7. Taylor diagram representing global statistics of CrIS (red), AIRS (blue), and joint AIRS+OMI (green) tropospheric column comparisons with sondes. The radial distance from the origin represents standard deviation, azimuthal angle represents the correlation coefficient (r), and the dashed arcs centered on the x-axis represent the root mean square error (RMSE). Markers represent datasets in each season (DJF: December/January/February, MAM: March/April/May, JJA: June/July/August, SON: September/October/November) and annually.

CrIS have similar seasonal patterns, with the smallest standard deviations in SON and the largest standard deviations in DJF. The seasonal spread of AIRS and CrIS is relatively small in the LT, likely because of their low sensitivity there, with most of the seasonal differences occurring in the UT. High standard deviations in the UT in DJF are caused by high ozone values (Fig. S11) occurring in the Northern high latitudes (Fig. S12). Figures S11-S12 are presented for AIRS, but similar ozone distributions are seen in the CrIS dataset. Joint AIRS+OMI does not include high ozone values in the Northern high latitudes due to the low joint AIRS+OMI data volume near the poles. Because the high ozone values are seen in both the satellite records and the sondes, they are likely to be real. Previous work used the Modern-Era Retrospective analysis for Research and Applications Version 2 (MERRA-2) global reanalysis model to show the prevalence of multiple tropopauses at high Northern latitudes (Manney et al., 2017). Multiple tropopauses allow for enhanced mixing between the troposphere and stratosphere, allowing for high stratospheric ozone concentrations to intrude into the troposphere and raise ozone concentrations. Manney et al. (2017) also found a high frequency of multiple tropopauses in high Southern latitudes in JJA, but we do not observe high ozone values in JJA in the TROPESS datasets due to the lack of coincident satellite and sonde measurements below 60°S.

Figure 7 also demonstrates that joint AIRS+OMI has the lowest correlation with sonde measurements in all vertical regimes, with the largest difference between joint AIRS+OMI and the other two satellite products in the LT. Joint AIRS+OMI has the highest sensitivity in the LT compared to AIRS and CrIS (see Fig. S4) because it includes UV radiances from OMI, which are sensitive to the entire ozone column, including the surface (Fu et al., 2018). While it is clear that there is a systematic bias in the LT TROPESS joint AIRS+OMI retrievals, the relative lack of error in AIRS and CrIS likely arises because their retrievals provide little information in the LT due to low sensitivity. Since the satellite operator is applied to the ozonesonde measurements, this lack of sensitivity results in both the satellite and sonde profiles largely reflecting the prior.

https://doi.org/10.5194/egusphere-2024-3701
Preprint. Discussion started: 19 December 2024

The satellite-sonde percent bias also varies across season (Tables 2, 3, 4). In the Northern Hemisphere, the bias is largest for all three products in DJF and smallest in JJA. This seasonal pattern is also reflected in the global values since most of the data falls in the Northern Hemisphere. In the Southern Hemisphere, the change in bias across seasons is less consistent, but this may be due to the lack of data in the Southern Hemisphere compared to the Northern Hemisphere. In all seasons, the CrIS and joint AIRS+OMI biases remain positive (except 30-60°S CrIS in DJF), whereas the AIRS bias is negative in the spring (MAM) and summer (JJA) in the Northern Hemisphere and spring (SON) in the Southern Hemisphere.

Seasonality is also seen in the trend in satellite-sonde bias and can explain why the trend in tropical percent bias is larger than that in other regions (Tables 2, 3, 4). There is little consistency in the seasonality of the bias trend across products, but it is noteworthy that the magnitudes of the seasonal bias trends are substantially larger for CrIS than the annual average trend in most regions. However, we caution against over-interpreting these results given the relatively short duration of the CrIS product. In many regions, the bias trend fluctuates between positive and negative values across seasons, resulting in a small annual average value for each product. But in the tropics (15°S-15°N), the bias trend is almost always positive, leading to the large trend in annual bias, particularly for CrIS (Fig. 5c). The satellite sensitivity to ozone tends to be largest in the tropics due to a combination of good thermal contrast, warm atmospheric temperatures, high insolation, and large ozone abundances. The variability in these parameters is also lowest in the tropics, so that the sensitivity is more likely to be consistent across seasons. The seasonal dependence of the biases and how they vary across region and for each satellite product are important to consider when quantifying a trend in only one season or month. For example, some studies quantify the tropospheric ozone trend in summer only, since summer typically has the highest ozone values and worst air pollution. The impact of time-dependent satellite error can be different in the summer season than for the annual average, so the error should be considered explicitly.

Table 2: Comparisons between CrIS and sonde tropospheric ozone columns: bias (%), trend (% decade^{-1}), error (standard error on the trend), and p-value on the trend. Data are separated by region and season (ann = annual, DJF = December/January/February, MAM = March/April/May, JJA = June/July/August, SON = September/October/November).

	CrIS									
	all dates	DJF	MAM	JJA	SON	all dates	DJF	MAM	JJA	SON
	60°-90°N					**30°-60°N**				
N	5252	1894	1393	915	1050	6905	2112	1664	1320	1809
bias	5.0	6.5	3.8	3.7	5.2	5.2	7.3	6.0	2.7	4.1
trend	2.1	-4.5	-3.6	0.75	4.2	0.63	-4.6	2.0	0.57	-0.61
error	2.0	5.4	5.6	4.6	5.1	3.7	5.5	5.8	6.1	3.7
p-value	0.30	0.42	0.52	0.87	0.42	0.86	0.42	0.74	0.93	0.87
	0°-30°N					**15°S-15°N**				
N	696	221	155	146	174	958	271	238	238	211
bias	3.7	4.7	4.2	4.1	1.2	5.1	5.0	6.0	4.5	4.0
trend	-3.1	3.3	-8.4	0.75	1.3	9.0	7.2	8.1	19.2	5.9

https://doi.org/10.5194/egusphere-2024-3701
Preprint. Discussion started: 19 December 2024

error	3.7	9.3	9.4	11.1	8.8	4.6	10.8	8.0	9.9	8.2
p-value	0.40	0.73	0.38	0.95	0.88	0.05	0.51	0.33	0.07	0.48
	0°-30°S					**30°-60°S**				
N	736	174	192	192	178	166	57	39	30	40
bias	4.5	5.8	3.7	5.3	3.7	4.34	-0.07	4.2	8.0	1.5
trend	4.9	-0.26	0.60	4.6	-12.5	-0.40	1.8	0.83	6.9	0.10
error	5.1	14.6	16.9	17.7	8.1	7.4	32.8	16.6	22.2	23.7
p-value	0.33	0.99	0.97	0.80	0.14	0.96	0.96	0.96	0.76	1.0
	global									
N	13755	4458	3443	2603	3251					
bias	4.8	7.1	5.1	3.0	4.1					
trend	0.21	-1.0	1.0	3.3	0.82					
error	3.6	2.6	4.3	3.5	3.9					
p-value	0.95	0.70	0.81	0.36	0.83					

Table 3: Same as Table 2 but for AIRS.

	AIRS									
	all dates	DJF	MAM	JJA	SON	all dates	DJF	MAM	JJA	SON
	60°-90°N					**30°-60°N**				
N	15847	5410	4131	2892	3414	20438	4683	5127	5296	5332
bias	0.80	2.4	-0.52	-1.1	2.5	0.35	2.8	-0.47	-1.1	0.63
trend	-1.4	-1.7	0.30	-1.3	-1.8	-0.82	-1.2	0.06	-0.26	-1.5
error	0.46	0.65	0.65	0.63	0.91	0.48	0.73	0.82	0.73	1.0
p-value	0.002	0.01	0.65	0.04	0.05	0.09	0.11	0.95	0.72	0.16
	0°-30°N					**15°S-15°N**				
N	1869	451	421	503	494	2697	629	597	763	708
bias	1.0	1.2	1.9	-0.01	0.97	0.79	1.6	1.5	0.20	-0.33
trend	0.06	-1.2	-0.34	-0.29	1.8	1.8	0.77	2.4	2.9	-0.01
error	0.99	1.4	1.9	2.0	1.6	0.64	0.89	1.0	1.4	1.8
p-value	0.95	0.37	0.86	0.89	0.29	0.01	0.39	0.02	0.04	0.99
	0°-30°S					**30°-60°S**				
N	2337	507	561	700	569	460	90	123	113	134
bias	0.98	1.2	2.1	0.85	0.24	1.8	1.9	0.07	7.4	-0.67

https://doi.org/10.5194/egusphere-2024-3701
Preprint. Discussion started: 19 December 2024

trend	1.2	-0.99	1.7	3.4	1.6	-1.4	-0.91	-7.1	-0.53	0.21
error	0.88	1.6	1.1	1.3	2.2	2.0	5.5	3.0	3.0	3.9
p-value	0.16	0.54	0.13	0.01	0.46	0.49	0.87	0.02	0.86	0.96
	global									
N	43735	11141	10363	9504	9943					
bias	0.36	2.2	-0.58	-1.2	1.0					
trend	-0.41	-0.58	0.19	0.17	-1.5					
error	0.57	0.58	0.68	0.54	0.87					
p-value	0.48	0.32	0.78	0.75	0.08					

Table 4: Same as Table 2 but for joint AIRS+OMI.

	Joint AIRS+OMI									
	all dates	DJF	MAM	JJA	SON	all dates	DJF	MAM	JJA	SON
	60°-90°N					**30°-60°N**				
N	968	93	247	410	218	7736	843	1804	3052	2037
bias	4.8	12.3	1.8	-0.17	7.3	8.4	13.5	9.1	2.8	9.5
trend	-1.3	-0.70	-2.2	1.7	-4.1	-1.1	0.97	-0.31	1.9	-2.6
error	2.0	4.3	3.9	2.7	2.7	1.4	1.3	1.3	0.99	1.8
p-value	0.53	0.87	0.57	0.54	0.14	0.44	0.47	0.82	0.06	0.15
	0°-30°N					**15°S-15°N**				
N	1269	222	283	416	348	1759	343	414	518	484
bias	7.5	11.7	5.9	6.3	5.5	6.3	10.3	8.8	5.4	4.2
trend	1.2	0.46	3.5	3.4	-1.7	3.9	8.0	2.9	5.9	-0.47
error	1.7	3.1	3.1	2.3	2.9	1.8	3.2	2.8	2.2	2.0
p-value	0.48	0.88	0.26	0.16	0.55	0.03	0.02	0.31	0.01	0.82
	0°-30°S					**30°-60°S**				
N	1549	288	397	435	429	200	63	67	28	42
bias	4.7	3.2	6.7	5.3	2.5	11.7	9.1	17.2	13.7	9.7
trend	1.9	3.6	1.8	1.7	-1.3	3.1	-0.16	7.4	-2.6	8.0
error	1.1	2.4	1.9	1.4	1.6	3.9	7.6	4.9	8.5	7.8
p-value	0.09	0.13	0.36	0.22	0.40	0.43	0.98	0.14	0.77	0.31
	global									
N	11722	1509	2798	4341	3074					

https://doi.org/10.5194/egusphere-2024-3701
Preprint. Discussion started: 19 December 2024

bias	6.8	11.5	7.2	3.4	7.4
trend	1.1	3.5	1.3	1.6	-1.5
error	1.0	1.2	1.4	0.89	1.2
p-value	0.27	0.00	0.36	0.08	0.22

4 Discussion

The focus of this study was to quantify satellite-sonde biases and their evolution with time. Nonetheless, we wish to emphasize that analyses must consider the time range used for trend detection, as a short time period may not provide enough information to detect trends with high precision. The CrIS dataset used here runs from late 2015 to mid 2021, and the uncertainty associated with the CrIS-sonde bias trends are high (Fig. 5 and Table S4), except in the tropics. Additionally, when the AIRS and joint AIRS+OMI datasets were cropped to the same time period as the CrIS record, the uncertainties on their bias trends were very large as well. The ability to quantify a drift with high precision depends on both the magnitude of the drift and its variability, with longer records required for smaller drifts and/or larger variability. It is critical that statistical methods such as block bootstrapping, which preserves the correlation between consecutive data points and therefore more accurately captures the variance of a time series than other methods, are used to assess statistical significance and determine whether a given record is long enough to detect a trend.

A substantial portion of the effort underlying this analysis was directed at establishing QC methods for the ozonesonde data (Sect. 2.2.2 and SI), and we wish to emphasize that this process can impact the outcome of validation (i.e., bias estimation) studies. While the ozonesonde data has low measurement uncertainty and has historically been used to represent "true" ozone values, some of the sonde data has poor quality or does not interact with the satellite operator in a physically realistic manner. In some cases, the ozonesonde concentrations are realistic in the stratosphere and not the troposphere, but the application of the satellite operator convolves the stratosphere and troposphere, making a profile unusable. Typically, studies that utilize sonde data develop their own QC methods, especially when working with satellite operators from a variety of products. As shown in Sect. S1, the QC methods and subset of sonde data used can impact the quantification of satellite-sonde bias. So studies that use different QC methods may not be comparing their datasets to the same "true" information.

Section S1 suggests important topics for users of sonde data to consider. First, sonde users should investigate the impact that QC methods have on their results. This is true for ozone trend studies and studies that quantify instrument bias using sonde data as the ground truth dataset. As illustrated in Sect. S1, the trend in satellite-sonde bias varied by an order of magnitude and sign, ranging from -0.24 to 0.21% decade^{-1} (Table S2), depending on the QC method selected. The median bias and trend in bias stabilized when certain methods were employed (Tables S2, S3), suggesting that some methods provide more realistic and stable metrics than other methods. In addition to considering the stability of the bias, users should also consider the vertical profile of the bias, as different methods resulted in more or less physically realistic profiles at different levels (Fig. S2). Second, applying QC methods to the original sonde data may not be sufficient to completely QC the data when working with satellite operators, since the application of the satellite operator can introduce physically unrealistic behavior. On the

https://doi.org/10.5194/egusphere-2024-3701
Preprint. Discussion started: 19 December 2024

other hand, studies focusing solely on the sonde data will need to apply QC methods to only the raw sonde data. Because observational operators vary by instrument and in both space and time, we cannot state that our optimal QC method is tailored for use with other datasets. Third, QC methods can use historical or modeled data (e.g., from satellites, ground-based systems, and global models) to provide physically realistic concentration limits. In this study, the MLS climatology provided physical reasoning for the screening of unrealistic profiles in the stratosphere. Future work could investigate the curation of a long-term ozone climatology in the troposphere for use in QC screening. In summary, we recommend that future studies utilizing sonde data consider the impact of QC methods and how the chosen subset of sonde data compares to existing studies.

Another important consideration is whether the sonde sites are truly representative of global ozone and can be used to accurately quantify global satellite-sonde bias. The ozonesonde sites are mostly grouped in the Northern mid to high latitudes (Fig. 2 and 4), with little coverage in the Southern Hemisphere and most continents apart from Europe and North America. Chang et al. (2024) showed that the sampling of hourly ozone measurements at the Mauna Loa Observatory could drastically impact the accuracy of long-term trend detection. This work also showed that small-scale meteorological variability affects the trends. These results were given at one site, but demonstrate the impact of localized phenomena on trend detection from *in situ* observations, suggesting that individual sites may not be representative of regional trends if their temporal sampling or meteorology are inconsistent. Future work will focus on assessing the degree to which the satellite-sonde intercomparisons shown here are representative of broader changes in bias using a tropospheric chemistry reanalysis (Miyazaki et al., 2020a, b).

This study suggests important considerations for the TOAR community when performing validation or trend analyses that compare multiple satellite products. In addition to the factors presented above (i.e., time range, sonde QC methods, and geographic representativeness), the sensitivity of the instruments, the *a priori* information, and other factors in the retrievals algorithms can impact trend detection. The TROPESS products used in this study used a consistent retrieval algorithm that used the same *a priori* information, tropopause definitions, and vertical grids for all products. In an ideal scenario, all satellite products used in intercomparison studies would also use these consistent variables. Importantly, they would also all be compared to ozonesonde or other *in situ* data using the same filtering criteria, rather than comparing to a variable ground truth dataset. While this would require significant effort to accomplish the large number of available satellite products, these steps would allow for the most accurate comparison of satellite measurement error and quantification of tropospheric ozone.

5 Conclusions

This study compares long-term tropospheric ozone records from three satellite products and ozonesondes. The CrIS, AIRS, and joint AIRS+OMI products are those developed by the TROPESS project, which uses a common retrieval algorithm and consistent *a priori* information to provide ozone profiles with quantified uncertainties. A major goal of the TOAR-II project is to quantify long-term trends in tropospheric ozone, which requires understanding to what degree the observed trend is attributable to non-physical properties such as instrument measurement drift. The TROPESS products were compared to global ozonesonde data and the magnitude of the trends in satellite-sonde bias for the three products (CrIS: 0.21 ± 3.6% decade^{-1}, AIRS: -0.41 ± 0.57% decade^{-1}, and joint AIRS+OMI: 1.1 ± 1.0% decade^{-1}) are approximately one order of magnitude less than the

https://doi.org/10.5194/egusphere-2024-3701
Preprint. Discussion started: 19 December 2024

reported range in tropospheric ozone trends (-7.1% decade^{-1} to 9.5% decade^{-1}). This work suggests that the three TROPESS products can be used to accurately detect global tropospheric ozone trends. Future work to quantify trends in specific regions or seasons, however, should consider the impact that more localized satellite measurement drift has on trend detection. While the measurement drift could not be quantified with high precision in many regions or seasons (Fig. 5 and Tables 2 - 4), the tropics displayed unique behavior with consistently positive and non-zero drift in all instruments.

Data availability. TROPESS satellite data is available at the NASA GES-DISC database (https://disc.gsfc.nasa.gov/datasets?page=1&project=TROPESS). TROPESS datasets with the complete set of variables are available by contacting the corresponding author. Information about acquiring harmonized ozonesonde data is available at https://hegiftom.meteo.be/datasets/ozonesondes (HEGIFTOM, 2024).

Author contributions. All authors conceptualized the project, provided technical guidance, and contributed to writing and editing the article. EAP performed data analysis and led the writing and editing of the article.

Competing interests. The authors have no competing interests to declare.

Acknowledgements. The authors would like to thank Roeland Van Malderan for providing and assisting with the ozonesonde data. We also thank Frank Werner for processing the MLS climatology data. The research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration (80NM0018D0004). We also acknowledge the support of the NASA Atmospheric Composition: Aura Science Team Program (19-AURAST19-0044), Earth Science U.S. Participating Investigator program (22-EUSPI22-0005), ACMAP (22-ACMAP22-0013), and the NASA TROPESS project.

https://doi.org/10.5194/egusphere-2024-3701
Preprint. Discussion started: 19 December 2024

References

AIRS: AIRS/Aqua L1B Infrared (IR) geolocated and calibrated radiances V005, https://doi.org/10.5067/YZEXEVN4JGGJ, 2007.

AIRS: Aqua/AIRS L2 Near Real Time (NRT) Standard Physical Retrieval (AIRS-only) V7.0, https://doi.org/10.5067/RAEHAOH4VZM5, 2019.

Aumann, H., Chahine, M., Gautier, C., Goldberg, M., Kalnay, E., McMillin, L., Revercomb, H., Rosenkranz, P., Smith, W., Staelin, D., Strow, L., and Susskind, J.: AIRS/AMSU/HSB on the Aqua mission: design, science objectives, data products, and processing systems, IEEE Transactions on Geoscience and Remote Sensing, 41, 253–264, https://doi.org/10.1109/TGRS.2002.808356, conference Name: IEEE Transactions on Geoscience and Remote Sensing, 2003.

Barnet, C. D., Divakarla, M., Gambacorta, A., Iturbide-Sanchez, F., Nalli, N. R., Pryor, K., Tan, C., Wang, T., Warner, J., Zhang, K., and Zhu, T.: NOAA Unique Combined Atmospheric Processing System (NUCAPS): Algorithm Theoretical Basis Document, Tech. Rep. Version 3.1, NOAA NESDIS STAR, https://www.star.nesdis.noaa.gov/jpss/documents/ATBD/ATBD_NUCAPS_v3.1.pdf, 2021.

Beer, R.: TES on the aura mission: scientific objectives, measurements, and analysis overview, IEEE Transactions on Geoscience and Remote Sensing, 44, 1102–1105, https://doi.org/10.1109/TGRS.2005.863716, conference Name: IEEE Transactions on Geoscience and Remote Sensing, 2006.

Bowman, K., Rodgers, C., Kulawik, S., Worden, J., Sarkissian, E., Osterman, G., Steck, T., Lou, M., Eldering, A., Shephard, M., Worden, H., Lampel, M., Clough, S., Brown, P., Rinsland, C., Gunson, M., and Beer, R.: Tropospheric emission spectrometer: retrieval method and error analysis, IEEE Transactions on Geoscience and Remote Sensing, 44, 1297–1307, https://doi.org/10.1109/TGRS.2006.871234, conference Name: IEEE Transactions on Geoscience and Remote Sensing, 2006.

Boynard, A., Hurtmans, D., Garane, K., Goutail, F., Hadji-Lazaro, J., Koukouli, M. E., Wespes, C., Vigouroux, C., Keppens, A., Pommereau, J.-P., Pazmino, A., Balis, D., Loyola, D., Valks, P., Sussmann, R., Smale, D., Coheur, P.-F., and Clerbaux, C.: Validation of the IASI FORLI/EUMETSAT ozone products using satellite (GOME-2), ground-based (Brewer–Dobson, SAOZ, FTIR) and ozonesonde measurements, Atmospheric Measurement Techniques, 11, 5125–5152, https://doi.org/10.5194/amt-11-5125-2018, publisher: Copernicus GmbH, 2018.

Brewer, A. W., Milford, J. R., and Dobson, G. M. B.: The Oxford-Kew ozone sonde, Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences, 256, 470–495, https://doi.org/10.1098/rspa.1960.0120, publisher: Royal Society, 1997.

Chang, K.-L., Schultz, M. G., Koren, G., and Selke, N.: Guidance note on best statistical practices for TOAR analyses, https://igacproject.org/sites/default/files/2023-04/STAT_recommendations_TOAR_analyses_0.pdf, 2023.

Chang, K.-L., Cooper, O. R., Gaudel, A., Petropavlovskikh, I., Effertz, P., Morris, G., and McDonald, B. C.: Technical note: Challenges in detecting free tropospheric ozone trends in a sparsely sampled environment, Atmospheric Chemistry and Physics, 24, 6197–6218, https://doi.org/10.5194/acp-24-6197-2024, publisher: Copernicus GmbH, 2024.

Elshorbany, Y., Ziemke, J. R., Strode, S., Petetin, H., Miyazaki, K., De Smedt, I., Pickering, K., Seguel, R. J., Worden, H., Emmerichs, T., Taraborrelli, D., Cazorla, M., Fadnavis, S., Buchholz, R. R., Gaubert, B., Rojas, N. Y., Nogueira, T., Salameh, T., and Huang, M.: Tropospheric ozone precursors: global and regional distributions, trends, and variability, Atmospheric Chemistry and Physics, 24, 12 225–12 257, https://doi.org/10.5194/acp-24-12225-2024, publisher: Copernicus GmbH, 2024.

Fadnavis, S., Elshorbany, Y., Ziemke, J., Barret, B., Rap, A., Chandran, P. R. S., Pope, R., Sagar, V., Taraborrelli, D., Le Flochmoen, E., Cuesta, J., Wespes, C., Boersma, F., Glissenaar, I., De Smedt, I., Van Roozendael, M., Petetin, H., and Anglou, I.: Influence of nitrogen oxides and volatile organic compounds emission changes on tropospheric ozone variability, trends and radiative effect, EGUsphere, pp. 1–51, https://doi.org/10.5194/egusphere-2024-3050, publisher: Copernicus GmbH, 2024.

https://doi.org/10.5194/egusphere-2024-3701
Preprint. Discussion started: 19 December 2024

Froidevaux, L., Kinnison, D. E., Gaubert, B., Schwartz, M. J., Livesey, N. J., Read, W. G., Bardeen, C. G., Ziemke, J. R., and Fuller, R. A.: Tropical upper tropospheric trends in ozone and carbon monoxide (2005–2020): observational and model results, EGUsphere, pp. 1–59, https://doi.org/10.5194/egusphere-2024-525, publisher: Copernicus GmbH, 2024.

Fu, D., Worden, J. R., Liu, X., Kulawik, S. S., Bowman, K. W., and Natraj, V.: Characterization of ozone profiles derived from Aura TES and OMI radiances, Atmospheric Chemistry and Physics, 13, 3445–3462, https://doi.org/10.5194/acp-13-3445-2013, 2013.

Fu, D., Bowman, K. W., Worden, H. M., Natraj, V., Worden, J. R., Yu, S., Veefkind, P., Aben, I., Landgraf, J., Strow, L., and Han, Y.: High-resolution tropospheric carbon monoxide profiles retrieved from CrIS and TROPOMI, Atmospheric Measurement Techniques, 9, 2567–2579, https://doi.org/10.5194/amt-9-2567-2016, 2016.

Fu, D., Kulawik, S. S., Miyazaki, K., Bowman, K. W., Worden, J. R., Eldering, A., Livesey, N. J., Teixeira, J., Irion, F. W., Herman, R. L., Osterman, G. B., Liu, X., Levelt, P. F., Thompson, A. M., and Luo, M.: Retrievals of tropospheric ozone profiles from the synergism of AIRS and OMI: methodology and validation, Atmospheric Measurement Techniques, 11, 5587–5605, https://doi.org/10.5194/amt-11-5587-2018, 2018.

Gaudel, A., Cooper, O. R., Ancellet, G., Barret, B., Boynard, A., Burrows, J. P., Clerbaux, C., Coheur, P.-F., Cuesta, J., Cuevas, E., Doniki, S., Dufour, G., Ebojie, F., Foret, G., Garcia, O., Granados-Muñoz, M. J., Hannigan, J. W., Hase, F., Hassler, B., Huang, G., Hurtmans, D., Jaffe, D., Jones, N., Kalabokas, P., Kerridge, B., Kulawik, S., Latter, B., Leblanc, T., Le Flochmoën, E., Lin, W., Liu, J., Liu, X., Mahieu, E., McClure-Begley, A., Neu, J. L., Osman, M., Palm, M., Petetin, H., Petropavlovskikh, I., Querel, R., Rahpoe, N., Rozanov, A., Schultz, M. G., Schwab, J., Siddans, R., Smale, D., Steinbacher, M., Tanimoto, H., Tarasick, D. W., Thouret, V., Thompson, A. M., Trickl, T., Weatherhead, E., Wespes, C., Worden, H. M., Vigouroux, C., Xu, X., Zeng, G., and Ziemke, J.: Tropospheric Ozone Assessment Report: Present-day distribution and trends of tropospheric ozone relevant to climate and global atmospheric chemistry model evaluation, Elementa: Science of the Anthropocene, 6, 39, https://doi.org/10.1525/elementa.291, 2018.

Gaudel, A., Bourgeois, I., Li, M., Chang, K.-L., Ziemke, J., Sauvage, B., Stauffer, R. M., Thompson, A. M., Kollonige, D. E., Smith, N., Hubert, D., Keppens, A., Cuesta, J., Heue, K.-P., Veefkind, P., Aikin, K., Peischl, J., Thompson, C. R., Ryerson, T. B., Frost, G. J., McDonald, B. C., and Cooper, O. R.: Tropical tropospheric ozone distribution and trends from in situ and satellite data, Atmospheric Chemistry and Physics, 24, 9975–10 000, https://doi.org/10.5194/acp-24-9975-2024, publisher: Copernicus GmbH, 2024.

GMAO: GEOS FP-IT data, https://gmao.gsfc.nasa.gov/GMAO_products/NRT_products.php, 2024.

Han, Y.: CrIS Full Spectral Resolution SDR and S-NPP/JPSS-1 CrIS Performance Status, https://www.star.nesdis.noaa.gov/jpss/documents/meetings/2015/ITSC-XX/Yong_Han_2015ITSC-20_1p07.pdf, 2015.

Han, Y., Revercomb, H., Cromp, M., Gu, D., Johnson, D., Mooney, D., Scott, D., Strow, L., Bingham, G., Borg, L., Chen, Y., DeSlover, D., Esplin, M., Hagan, D., Jin, X., Knuteson, R., Motteler, H., Predina, J., Suwinski, L., Taylor, J., Tobin, D., Tremblay, D., Wang, C., Wang, L., Wang, L., and Zavyalov, V.: Suomi NPP CrIS measurements, sensor data record algorithm, calibration and validation activities, and record data quality, Journal of Geophysical Research: Atmospheres, 118, 12,734–12,748, https://doi.org/10.1002/2013JD020344, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/2013JD020344, 2013.

HEGIFTOM: Ozonesondes: Harmonization and Evaluation of Ground-based Instruments for Free Tropospheric Ozone Measurements., https://hegiftom.meteo.be/datasets/ozonesondes, 2024.

Heue, K.-P., Coldewey-Egbers, M., Delcloo, A., Lerot, C., Loyola, D., Valks, P., and van Roozendael, M.: Trends of tropical tropospheric ozone from 20 years of European satellite measurements and perspectives for the Sentinel-5 Precursor, Atmospheric Measurement Techniques, 9, 5037–5051, https://doi.org/10.5194/amt-9-5037-2016, publisher: Copernicus GmbH, 2016.

https://doi.org/10.5194/egusphere-2024-3701
Preprint. Discussion started: 19 December 2024

Huang, G., Liu, X., Chance, K., Yang, K., Bhartia, P. K., Cai, Z., Allaart, M., Ancellet, G., Calpini, B., Coetzee, G. J. R., Cuevas-Agulló, E., Cupeiro, M., De Backer, H., Dubey, M. K., Fuelberg, H. E., Fujiwara, M., Godin-Beekmann, S., Hall, T. J., Johnson, B., Joseph, E., Kivi, R., Kois, B., Komala, N., König-Langlo, G., Laneve, G., Leblanc, T., Marchand, M., Minschwaner, K. R., Morris, G., Newchurch, M. J., Ogino, S.-Y., Ohkawara, N., Piters, A. J. M., Posny, F., Querel, R., Scheele, R., Schmidlin, F. J., Schnell, R. C., Schrems, O., Selkirk, H., Shiotani, M., Skrivánková, P., Stübi, R., Taha, G., Tarasick, D. W., Thompson, A. M., Thouret, V., Tully, M. B., Van Malderen, R., Vömel, H., von der Gathen, P., Witte, J. C., and Yela, M.: Validation of 10-year SAO OMI Ozone Profile (PROFOZ) product using ozonesonde observations, Atmospheric Measurement Techniques, 10, 2455–2475, https://doi.org/10.5194/amt-10-2455-2017, publisher: Copernicus GmbH, 2017.

Iturbide-Sanchez, F., Strow, L., Tobin, D., Chen, Y., Tremblay, D., Knuteson, R. O., Johnson, D. G., Buttles, C., Suwinski, L., Thomas, B. P., Rivera, A. R., Lynch, E., Zhang, K., Wang, Z., Porter, W. D., Jin, X., Predina, J. P., Eresmaa, R. I., Collard, A., Ruston, B., Jung, J. A., Barnet, C. D., Beierle, P. J., Yan, B., Mooney, D., and Revercomb, H.: Recalibration and Assessment of the SNPP CrIS Instrument: A Successful History of Restoration After Midwave Infrared Band Anomaly, IEEE Transactions on Geoscience and Remote Sensing, 60, 1–21, https://doi.org/10.1109/TGRS.2021.3112400, conference Name: IEEE Transactions on Geoscience and Remote Sensing, 2022.

Jiang, Y. B., Froidevaux, L., Lambert, A., Livesey, N. J., Read, W. G., Waters, J. W., Bojkov, B., Leblanc, T., McDermid, I. S., Godin-Beekmann, S., Filipiak, M. J., Harwood, R. S., Fuller, R. A., Daffer, W. H., Drouin, B. J., Cofield, R. E., Cuddy, D. T., Jarnot, R. F., Knosp, B. W., Perun, V. S., Schwartz, M. J., Snyder, W. V., Stek, P. C., Thurstans, R. P., Wagner, P. A., Allaart, M., Andersen, S. B., Bodeker, G., Calpini, B., Claude, H., Coetzee, G., Davies, J., De Backer, H., Dier, H., Fujiwara, M., Johnson, B., Kelder, H., Leme, N. P., König-Langlo, G., Kyro, E., Laneve, G., Fook, L. S., Merrill, J., Morris, G., Newchurch, M., Oltmans, S., Parrondos, M. C., Posny, F., Schmidlin, F., Skrivankova, P., Stubi, R., Tarasick, D., Thompson, A., Thouret, V., Viatte, P., Vömel, H., von Der Gathen, P., Yela, M., and Zablocki, G.: Validation of Aura Microwave Limb Sounder Ozone by ozonesonde and lidar measurements, Journal of Geophysical Research: Atmospheres, 112, https://doi.org/10.1029/2007JD008776, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2007JD008776, 2007.

Jones, D. B. A., Bowman, K. W., Palmer, P. I., Worden, J. R., Jacob, D. J., Hoffman, R. N., Bey, I., and Yantosca, R. M.: Potential of observations from the Tropospheric Emission Spectrometer to constrain continental sources of carbon monoxide, Journal of Geophysical Research: Atmospheres, 108, https://doi.org/10.1029/2003JD003702, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2003JD003702, 2003.

Komhyr, W. D.: Electrochemical concentration cells for gas analysis, Annales Geophysicae, pp. 203–210, 1969.

Komhyr, W. D. and Harris, T. B.: Development of an ECC Ozonesonde, Tech. Rep. ERL 200-APCL 18, 1971.

Levelt, P., van den Oord, G., Dobber, M., Malkki, A., Visser, H., Vries, J. d., Stammes, P., Lundell, J., and Saari, H.: The ozone monitoring instrument, IEEE Transactions on Geoscience and Remote Sensing, 44, 1093–1101, https://doi.org/10.1109/TGRS.2006.872333, conference Name: IEEE Transactions on Geoscience and Remote Sensing, 2006.

Levelt, P. F., Joiner, J., Tamminen, J., Veefkind, J. P., Bhartia, P. K., Stein Zweers, D. C., Duncan, B. N., Streets, D. G., Eskes, H., van der A, R., McLinden, C., Fioletov, V., Carn, S., de Laat, J., DeLand, M., Marchenko, S., McPeters, R., Ziemke, J., Fu, D., Liu, X., Pickering, K., Apituley, A., González Abad, G., Arola, A., Boersma, F., Chan Miller, C., Chance, K., de Graaf, M., Hakkarainen, J., Hassinen, S., Ialongo, I., Kleipool, Q., Krotkov, N., Li, C., Lamsal, L., Newman, P., Nowlan, C., Suleiman, R., Tilstra, L. G., Torres, O., Wang, H., and Wargan, K.: The Ozone Monitoring Instrument: overview of 14 years in space, Atmospheric Chemistry and Physics, 18, 5699–5745, https://doi.org/10.5194/acp-18-5699-2018, publisher: Copernicus GmbH, 2018.

https://doi.org/10.5194/egusphere-2024-3701
Preprint. Discussion started: 19 December 2024

Leventidou, E., Weber, M., Eichmann, K.-U., Burrows, J. P., Heue, K.-P., Thompson, A. M., and Johnson, B. J.: Harmonisation and trends of 20-year tropical tropospheric ozone data, Atmospheric Chemistry and Physics, 18, 9189–9205, https://doi.org/10.5194/acp-18-9189-2018, publisher: Copernicus GmbH, 2018.

Liu, X., Bhartia, P. K., Chance, K., Froidevaux, L., Spurr, R. J. D., and Kurosu, T. P.: Validation of Ozone Monitoring Instrument (OMI) ozone profiles and stratospheric ozone columns with Microwave Limb Sounder (MLS) measurements, Atmospheric Chemistry and Physics, 10, 2539–2549, https://doi.org/10.5194/acp-10-2539-2010, publisher: Copernicus GmbH, 2010a.

Liu, X., Bhartia, P. K., Chance, K., Spurr, R. J. D., and Kurosu, T. P.: Ozone profile retrievals from the Ozone Monitoring Instrument, Atmospheric Chemistry and Physics, 10, 2521–2537, https://doi.org/10.5194/acp-10-2521-2010, publisher: Copernicus GmbH, 2010b.

Livesey, N. J., Read, W. G., Wagner, P. A., Froidevaux, L., Santee, M. L., Schwartz, M. J., Lambert, A., Millan Valle, L. F., Pumphrey, H. C., Manney, G. L., Fuller, R. A., Jarnot, R. F., Knosp, B. W., and Lay, R. R.: Version 5.0x Level 2 and 3 data quality and description document, Tech. Rep. JPL D-105336 Rev. B, Jet Propulsion Laboratory, California Institute of Technology, https://mls.jpl.nasa.gov/data/v5-0_data_quality_document.pdf, 2022.

Lucchesi, R.: File Specification for GEOS-5 FP-IT (Forward Processing for Instrument Teams), Tech. Rep. GMAO Office Note No. 2 (Version 1.3), https://gmao.gsfc.nasa.gov/pubs/docs/Lucchesi865.pdf, 2015.

Malina, E., Bowman, K. W., Kantchev, V., Kuai, L., Kurosu, T. P., Miyazaki, K., Natraj, V., Osterman, G. B., Oyafuso, F., and Thill, M. D.: Joint spectral retrievals of ozone with Suomi NPP CrIS augmented by S5P/TROPOMI, Atmospheric Measurement Techniques, 17, 5341–5371, https://doi.org/10.5194/amt-17-5341-2024, publisher: Copernicus GmbH, 2024.

Manney, G. L., Hegglin, M. I., Lawrence, Z. D., Wargan, K., Millán, L. F., Schwartz, M. J., Santee, M. L., Lambert, A., Pawson, S., Knosp, B. W., Fuller, R. A., and Daffer, W. H.: Reanalysis comparisons of upper tropospheric–lower stratospheric jets and multiple tropopauses, Atmospheric Chemistry and Physics, 17, 11 541–11 566, https://doi.org/10.5194/acp-17-11541-2017, publisher: Copernicus GmbH, 2017.

Miles, G. M., Siddans, R., Kerridge, B. J., Latter, B. G., and Richards, N. a. D.: Tropospheric ozone and ozone profiles retrieved from GOME-2 and their validation, Atmospheric Measurement Techniques, 8, 385–398, https://doi.org/10.5194/amt-8-385-2015, publisher: Copernicus GmbH, 2015.

Miyazaki, K., Eskes, H., Sudo, K., Boersma, K. F., Bowman, K., and Kanaya, Y.: Decadal changes in global surface NO_x emissions from multi-constituent satellite data assimilation, Atmospheric Chemistry and Physics, 17, 807–837, https://doi.org/10.5194/acp-17-807-2017, publisher: Copernicus GmbH, 2017.

Miyazaki, K., Sekiya, T., Fu, D., Bowman, K. W., Kulawik, S. S., Sudo, K., Walker, T., Kanaya, Y., Takigawa, M., Ogochi, K., Eskes, H., Boersma, K. F., Thompson, A. M., Gaubert, B., Barre, J., and Emmons, L. K.: Balance of Emission and Dynamical Controls on Ozone During the Korea-United States Air Quality Campaign From Multiconstituent Satellite Data Assimilation, Journal of Geophysical Research: Atmospheres, 124, 387–413, https://doi.org/10.1029/2018JD028912, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2018JD028912, 2019.

Miyazaki, K., Bowman, K., Sekiya, T., Eskes, H., Boersma, F., Worden, H., Livesey, N., Payne, V. H., Sudo, K., Kanaya, Y., Takigawa, M., and Ogochi, K.: Updated tropospheric chemistry reanalysis and emission estimates, TCR-2, for 2005–2018, Earth System Science Data, 12, 2223–2259, https://doi.org/10.5194/essd-12-2223-2020, publisher: Copernicus GmbH, 2020a.

Miyazaki, K., Bowman, K. W., Yumimoto, K., Walker, T., and Sudo, K.: Evaluation of a multi-model, multi-constituent assimilation framework for tropospheric chemical reanalysis, Atmospheric Chemistry and Physics, 20, 931–967, https://doi.org/10.5194/acp-20-931-2020, publisher: Copernicus GmbH, 2020b.

https://doi.org/10.5194/egusphere-2024-3701
Preprint. Discussion started: 19 December 2024

Miyazaki, K., Bowman, K., Sekiya, T., Takigawa, M., Neu, J. L., Sudo, K., Osterman, G., and Eskes, H.: Global tropospheric ozone responses to reduced NOx emissions linked to the COVID-19 worldwide lockdowns, Science Advances, 7, eabf7460, https://doi.org/10.1126/sciadv.abf7460, publisher: American Association for the Advancement of Science, 2021.

Molod, A., Takacs, L., Suarez, M., Bacmeister, J., Song, I.-S., and Eichmann, A.: The GEOS-5 Atmospheric General Circulation Model: Mean Climate and Development from MERRA to Fortuna, Tech. Rep. NASA/TM–2012-104606/Vol 28, NASA Goddard Space Flight Center, Greenbelt, Maryland, 2012.

NASA: TROPESS, https://tes.jpl.nasa.gov/tropess/, https://tes.jpl.nasa.gov/tropess/, 2024.

Nassar, R., Logan, J. A., Worden, H. M., Megretskaia, I. A., Bowman, K. W., Osterman, G. B., Thompson, A. M., Tarasick, D. W., Austin, S., Claude, H., Dubey, M. K., Hocking, W. K., Johnson, B. J., Joseph, E., Merrill, J., Morris, G. A., Newchurch, M., Oltmans, S. J., Posny, F., Schmidlin, F. J., Vömel, H., Whiteman, D. N., and Witte, J. C.: Validation of Tropospheric Emission Spectrometer (TES) nadir ozone profiles using ozonesonde measurements, Journal of Geophysical Research: Atmospheres, 113, https://doi.org/10.1029/2007JD008819, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2007JD008819, 2008.

NOAA: CrIS on SNPP switch to side 1 on July 12, 2021, https://www.ospo.noaa.gov/data/messages/2021/MSG188150401.html, 2021.

Osterman, G. B., Kulawik, S. S., Worden, H. M., Richards, N. a. D., Fisher, B. M., Eldering, A., Shephard, M. W., Froidevaux, L., Labow, G., Luo, M., Herman, R. L., Bowman, K. W., and Thompson, A. M.: Validation of Tropospheric Emission Spectrometer (TES) measurements of the total, stratospheric, and tropospheric column abundance of ozone, Journal of Geophysical Research: Atmospheres, 113, https://doi.org/10.1029/2007JD008801, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2007JD008801, 2008.

Parrington, M., Jones, D. B. A., Bowman, K. W., Thompson, A. M., Tarasick, D. W., Merrill, J., Oltmans, S. J., Leblanc, T., Witte, J. C., and Millet, D. B.: Impact of the assimilation of ozone from the Tropospheric Emission Spectrometer on surface ozone across North America, Geophysical Research Letters, 36, https://doi.org/10.1029/2008GL036935, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2008GL036935, 2009.

Pope, R. J., Kerridge, B. J., Siddans, R., Latter, B. G., Chipperfield, M. P., Feng, W., Pimlott, M. A., Dhomse, S. S., Retscher, C., and Rigby, R.: Investigation of spatial and temporal variability in lower tropospheric ozone from RAL Space UV–Vis satellite products, Atmospheric Chemistry and Physics, 23, 14 933–14 947, https://doi.org/10.5194/acp-23-14933-2023, publisher: Copernicus GmbH, 2023.

Pope, R. J., O'Connor, F. M., Dalvi, M., Kerridge, B. J., Siddans, R., Latter, B. G., Barret, B., Le Flochmoen, E., Boynard, A., Chipperfield, M. P., Feng, W., Pimlott, M. A., Dhomse, S. S., Retscher, C., Wespes, C., and Rigby, R.: Investigation of the impact of satellite vertical sensitivity on long-term retrieved lower-tropospheric ozone trends, Atmospheric Chemistry and Physics, 24, 9177–9195, https://doi.org/10.5194/acp-24-9177-2024, publisher: Copernicus GmbH, 2024.

Revercomb, H. and Strow, L.: Suomi NPP CrIS Level 1B Full Spectral Resolution V2, https://doi.org/10.5067/9NPOTPIPLMAW, 2018.

Schenkeveld, V. M. E., Jaross, G., Marchenko, S., Haffner, D., Kleipool, Q. L., Rozemeijer, N. C., Veefkind, J. P., and Levelt, P. F.: In-flight performance of the Ozone Monitoring Instrument, Atmospheric Measurement Techniques, 10, 1957–1986, https://doi.org/10.5194/amt-10-1957-2017, publisher: Copernicus GmbH, 2017.

Smith, N. and Barnet, C. D.: CLIMCAPS observing capability for temperature, moisture, and trace gases from AIRS/AMSU and CrIS/ATMS, Atmospheric Measurement Techniques, 13, 4437–4459, https://doi.org/10.5194/amt-13-4437-2020, publisher: Copernicus GmbH, 2020.

Steinbrecht, W., Schwarz, R., and Claude, H.: New Pump Correction for the Brewer–Mast Ozone Sonde: Determination from Experiment and Instrument Intercomparisons, https://doi.org/https://doi.org/10.1175/1520-0426(1998)015<0144:NPCFTB>2.0.CO;2, section: Journal of Atmospheric and Oceanic Technology, 1998.

https://doi.org/10.5194/egusphere-2024-3701
Preprint. Discussion started: 19 December 2024

Stephens, G. L., Vane, D. G., Boain, R. J., Mace, G. G., Sassen, K., Wang, Z., Illingworth, A. J., O'connor, E. J., Rossow, W. B., Durden, S. L., Miller, S. D., Austin, R. T., Benedetti, A., and Mitrescu, C.: THE CLOUDSAT MISSION AND THE A-TRAIN, https://doi.org/10.1175/BAMS-83-12-1771, section: Bulletin of the American Meteorological Society, 2002.

Strow, L. L., Motteler, H., Tobin, D., Revercomb, H., Hannon, S., Buijs, H., Predina, J., Suwinski, L., and Glumb, R.: Spectral calibration and validation of the Cross-track Infrared Sounder on the Suomi NPP satellite, Journal of Geophysical Research: Atmospheres, 118, 12,486–12,496, https://doi.org/10.1002/2013JD020480, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/2013JD020480, 2013.

Susskind, J., Barnet, C., and Blaisdell, J.: Retrieval of atmospheric and surface parameters from AIRS/AMSU/HSB data in the presence of clouds, IEEE Transactions on Geoscience and Remote Sensing, 41, 390–409, https://doi.org/10.1109/TGRS.2002.808236, conference Name: IEEE Transactions on Geoscience and Remote Sensing, 2003.

Susskind, J., Blaisdell, J. M., and Iredell, L.: Improved methodology for surface and atmospheric soundings, error estimates, and quality control procedures: the atmospheric infrared sounder science team version-6 retrieval algorithm, Journal of Applied Remote Sensing, 8, 084 994, https://doi.org/10.1117/1.JRS.8.084994, publisher: SPIE, 2014.

Tarasick, D. W., Smit, H. G. J., Thompson, A. M., Morris, G. A., Witte, J. C., Davies, J., Nakano, T., Van Malderen, R., Stauffer, R. M., Johnson, B. J., Stübi, R., Oltmans, S. J., and Vömel, H.: Improving ECC Ozonesonde Data Quality: Assessment of Current Methods and Outstanding Issues, Earth and Space Science, 8, e2019EA000 914, https://doi.org/10.1029/2019EA000914, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2019EA000914, 2021.

Tobin, D., Revercomb, H., Knuteson, R., Taylor, J., Best, F., Borg, L., DeSlover, D., Martin, G., Buijs, H., Esplin, M., Glumb, R., Han, Y., Mooney, D., Predina, J., Strow, L., Suwinski, L., and Wang, L.: Suomi-NPP CrIS radiometric calibration uncertainty, Journal of Geophysical Research: Atmospheres, 118, 10,589–10,600, https://doi.org/10.1002/jgrd.50809, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/jgrd.50809, 2013.

Van Malderen, R.: TOAR-II HEGIFTOM: Description of homogenized ozonesonde free-tropospheric ozone time series, 2023.

Verstraeten, W. W., Boersma, K. F., Zörner, J., Allaart, M. a. F., Bowman, K. W., and Worden, J. R.: Validation of six years of TES tropospheric ozone retrievals with ozonesonde measurements: implications for spatial patterns and temporal stability in the bias, Atmospheric Measurement Techniques, 6, 1413–1423, https://doi.org/10.5194/amt-6-1413-2013, publisher: Copernicus GmbH, 2013.

Wang, L., Tremblay, D. A., Han, Y., Esplin, M., Hagan, D. E., Predina, J., Suwinski, L., Jin, X., and Chen, Y.: Geolocation assessment for CrIS sensor data records, Journal of Geophysical Research: Atmospheres, 118, 12,690–12,704, https://doi.org/10.1002/2013JD020376, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/2013JD020376, 2013.

Waters, J., Froidevaux, L., Harwood, R., Jarnot, R., Pickett, H., Read, W., Siegel, P., Cofield, R., Filipiak, M., Flower, D., Holden, J., Lau, G., Livesey, N., Manney, G., Pumphrey, H., Santee, M., Wu, D., Cuddy, D., Lay, R., Loo, M., Perun, V., Schwartz, M., Stek, P., Thurstans, R., Boyles, M., Chandra, K., Chavez, M., Chen, G.-S., Chudasama, B., Dodge, R., Fuller, R., Girard, M., Jiang, J., Jiang, Y., Knosp, B., LaBelle, R., Lam, J., Lee, K., Miller, D., Oswald, J., Patel, N., Pukala, D., Quintero, O., Scaff, D., Van Snyder, W., Tope, M., Wagner, P., and Walch, M.: The Earth observing system microwave limb sounder (EOS MLS) on the aura Satellite, IEEE Transactions on Geoscience and Remote Sensing, 44, 1075–1092, https://doi.org/10.1109/TGRS.2006.873771, conference Name: IEEE Transactions on Geoscience and Remote Sensing, 2006.

Wei, J. C., Pan, L. L., Maddy, E., Pittman, J. V., Divarkarla, M., Xiong, X., and Barnet, C.: Ozone Profile Retrieval from an Advanced Infrared Sounder: Experiments with Tropopause-Based Climatology and Optimal Estimation Approach, https://doi.org/10.1175/2010JTECHA1384.1, section: Journal of Atmospheric and Oceanic Technology, 2010.

https://doi.org/10.5194/egusphere-2024-3701
Preprint. Discussion started: 19 December 2024

Werner, F., Livesey, N. J., Millán, L. F., Read, W. G., Schwartz, M. J., Wagner, P. A., Daffer, W. H., Lambert, A., Tolstoff, S. N., and Santee, M. L.: Applying machine learning to improve the near-real-time products of the Aura Microwave Limb Sounder, Atmospheric Measurement Techniques, 16, 2733–2751, https://doi.org/10.5194/amt-16-2733-2023, publisher: Copernicus GmbH, 2023.

WMO: Meteorology - A three-dimensional science: Second session of the commission for aerology, WMO Bulletin, 6, 134–138, 1957.

WMO: Ozonesonde Measurement Principles and Best Operational Practices: ASOPOS 2.0 (Assessment of Standard Operating Procedures for Ozonesondes), GAW Report No. 268, World Meteorological Organization (WMO), Geneva, Switzerland, https://library.wmo.int/idurl/4/57720, 2021.

WMO/GAW: World Meteorological Organization-Global Atmosphere Watch Program (WMO-GAW)/World Ozone and Ultraviolet Radiation Data Centre (WOUDC), https://doi.org/doi:10.14287/10000001, a list of all contributors is available on the website, 2024.

Worden, H. M., Logan, J. A., Worden, J. R., Beer, R., Bowman, K., Clough, S. A., Eldering, A., Fisher, B. M., Gunson, M. R., Herman, R. L., Kulawik, S. S., Lampel, M. C., Luo, M., Megretskaia, I. A., Osterman, G. B., and Shephard, M. W.: Comparisons of Tropospheric Emission Spectrometer (TES) ozone profiles to ozonesondes: Methods and initial results, Journal of Geophysical Research: Atmospheres, 112, https://doi.org/10.1029/2006JD007258, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2006JD007258, 2007.

Wu, W., Liu, X., Lei, L., Xiong, X., Yang, Q., Yue, Q., Zhou, D. K., and Larar, A. M.: Single field-of-view sounder atmospheric product retrieval algorithm: establishing radiometric consistency for hyper-spectral sounder retrievals, Atmospheric Measurement Techniques, 16, 4807–4832, https://doi.org/10.5194/amt-16-4807-2023, publisher: Copernicus GmbH, 2023.

Ziemke, J. R., Oman, L. D., Strode, S. A., Douglass, A. R., Olsen, M. A., McPeters, R. D., Bhartia, P. K., Froidevaux, L., Labow, G. J., Witte, J. C., Thompson, A. M., Haffner, D. P., Kramarova, N. A., Frith, S. M., Huang, L.-K., Jaross, G. R., Seftor, C. J., Deland, M. T., and Taylor, S. L.: Trends in global tropospheric ozone inferred from a composite record of TOMS/OMI/MLS/OMPS satellite measurements and the MERRA-2 GMI simulation, Atmospheric Chemistry and Physics, 19, 3257–3269, https://doi.org/10.5194/acp-19-3257-2019, publisher: Copernicus GmbH, 2019.