AGU ADVANCING EARTH AND SPACE SCIENCES

AGU Advances

RESEARCH ARTICLE
10.1029/2025AV001651

Calibrating the Tropospheric Air and Ozone Mass

Michael J. Prather[1]

[1]Earth System Science Department, UC Irvine, Irvine, CA, USA

Peer Review The peer review history for this article is available as a PDF in the Supporting Information.

Key Points:

- For chemical budgets and climate trends, which differ greatly between stratosphere and troposphere, we need a clear boundary, the tropopause
- Age-of-air tracers cleanly define gradients across the tropopause and avoid failures of traditional methods relying on vertical gradients
- We calibrate age tracers at specific locations where traditional methods are well founded and then apply them globally to ozone and airmass

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

M. J. Prather,
mprather@uci.edu

Citation:

Prather, M. J. (2025). Calibrating the tropospheric air and ozone mass. *AGU Advances*, *6*, e2025AV001651. https://doi.org/10.1029/2025AV001651

Received 4 JAN 2025
Accepted 13 MAY 2025

Author Contributions:

Conceptualization: Michael J. Prather
Data curation: Michael J. Prather
Formal analysis: Michael J. Prather
Funding acquisition: Michael J. Prather
Investigation: Michael J. Prather
Methodology: Michael J. Prather
Project administration: Michael J. Prather
Resources: Michael J. Prather
Software: Michael J. Prather
Validation: Michael J. Prather
Visualization: Michael J. Prather
Writing – original draft: Michael J. Prather

Abstract We divide the atmosphere into distinct spheres based on their physical, chemical, and dynamical traits. In deriving chemical budgets and climate trends, which differ across spheres, we need clearly defined boundaries. Our primary spheres are the troposphere and stratosphere (~99.9% by mass), and the boundary between them is the tropopause. Every global climate-weather model has one or more methods to calculate the lapse rate tropopause, but these involve subjective choices and are known to fail near the sub-tropical jets and polar regions. Age-of-air tracers clock the effective time-distance from the tropopause, allowing unambiguous separation of stratosphere from troposphere in the chaotic jet regions. We apply a global model with synthetic tracer e90 (90-day e-folding), focusing on ozone and temperature structures about the tropopause using ozone sonde and satellite observations. We calibrate an observation-consistent tropopause for e90 using tropics-plus-midlatitudes and then apply it globally to calculate total tropospheric air-mass and tropopause ozone values. The tropopause mixing barrier for the current UCI CTM is identified by a transition in the vertical transport gradient to stratospheric values of 15 days km^{-1}, corresponding to an e90 tropopause at 81 ± 2 ppb with a global tropospheric air mass of 82.2 ± 0.3%. The best e90 tropopause based on sonde pressures is 70–80 ppb; but that for ozone is 80–90 ppb, implying that the CTM tropopause ozone values are too large. This approach of calibrating an age-of-air tropopause can be readily applied to other models and possibly used with observed age-of-air tracers like sulfur hexafluoride.

Plain Language Summary The atmosphere is split into distinct spheres based on chemical and dynamical traits with 99.9% of the mass in the two lowermost spheres, the troposphere and stratosphere. The troposphere is dynamically well mixed on time scales of weeks (vertical) to a year (interhemispheric),yet, paradoxically, the troposphere has large, almost chaotic chemical heterogeneity caused by the small-scale chemical sources (lightning, wildfires, pollution) and sinks (convection, rainfall). Because it mostly lacks these small-scale sources and sinks, the stratosphere contains smooth photochemically driven vertical gradients in chemical species. We need to separate these two spheres when assessing climate change because each forces climate and responds to it in their own way. This work assesses our ability to draw the boundary between stratosphere and troposphere, which is named the tropopause. We reaffirm the work of others that clearly shows the traditional method of defining the tropopause fails in some key regions of the atmosphere. We show here that an age-of-air tracer can be calibrated to match the traditional tropopause where it is well defined and then used to calculate the tropopause in the difficult regions about jet streams and polar regions. This novel approach can be used in global climate models and possibly observations.

1. Introduction

Accurately diagnosing air parcels as either tropospheric or stratospheric is important for our understanding of the atmospheric chemistry and physics that control greenhouse gases and air pollutants. For example, a tropospheric ozone (O_3) perturbation absorbs additional solar radiation that heats the troposphere and warms the climate system, but if that perturbation is stratospheric, then the same absorbed solar energy is prevented from reaching the troposphere, thus cooling the climate system (Skeie et al., 2020). While the tropopause is readily and consistently defined in certain locations and times by criteria based on dynamical quantities or trace gases, there are many parts of the globe where these methods fail. For example, we have observed the climate-driven expansion of the troposphere through the trend in tropopause pressure (Santer et al., 2003; Seidel & Randel, 2006; Xian & Homeyer, 2019). Unfortunately, this trend can only be detected in some regions because the tropopause frequently cannot be accurately defined around the sub-tropics or in polar regions (Reichler et al., 2003; Tinney et al., 2022). Thus, we have no ready measure of the global tropospheric air mass fraction (TAMF) or its trend.

Writing – review & editing: Michael J. Prather

This study of the tropopause continues the work of Prather et al. (2011, hence P2011) using the synthetic tracer e90 (uniform surface emissions, 90-day e-folding time) in the UC Irvine chemistry-transport model (UCI CTM) to characterize each air parcel as either tropospheric or stratospheric. Other researchers have found that in their models the tropopause threshold e90 value differed from the P2011 value of 90 ppb. This paper's goal is to combine a range of observations and traditional tropopause criteria with CTM simulations to calibrate the e90 tracer over a range of locations and seasons so that the e90 tracer can be used globally, even in the difficult transition regions about the jets. We can then readily calculate the tropospheric ozone and air mass as well as the ozone gradients across the tropopause in our models. This calibration method is tested here with a different set of meteorological data. It is designed for use by other models so that the community can develop more standard approaches to stratosphere-troposphere diagnostics.

2. Data

2.1. Global Chemistry Models

The problem in finding the tropopause (abbreviated "tpp" as an adjective) ozone value in a model is illustrated in Figure 1a, where we diagnose a single snapshot of a latitude-by-altitude meridional transect of O_3 abundance (nmol mol^{-1} of dry air, ppb) from the UCI CTM control simulations of Prather and Zhu (2024). The O_3 values are plotted (non-white rectangles) only for grid cells considered tropospheric based on the World Meteorological Organization (WMO) criteria of the lapse rate (LR = −dT/dz) being >2 K km^{-1} between that cell and the one above. Based on the original e90 tpp threshold of 90 ppb (P2011), we mark the upper e90 tropopause with a thick black line and the lower e90 tropopause a thin black line. We define the lower e90 tropopause as the top edge of the lowest tropospheric layer (above 4 km) with a stratospheric layer immediately above. We define the upper e90 tropopause as the top edge of the highest tropospheric layer. Most often, there is only a single tropopause in an air column. Near 35°S we see a stratospheric fold (white blocks) reaching down to 7 km altitude. The 6-km vertical discontinuity in the e90 tropopause at 18°S indicates a vertically aligned tropopause that is clearly identified here by horizontal gradients in O_3 and also e90 (not shown).

Above the sub-tropical jets and in the polar regions we find grid cells well separated from the lower tropopause with exceptionally high O_3 values > 200 ppb (brick red) being identified as tropospheric by the lapse rate tropopause (LRT) criterion. This example of LRT inconsistencies is typical (see Figures 3 and 4 of Hoffmann & Spang, 2022). In terms of where LRT uncertainties are greatest, these locations (jets and poles) are consistent with the thorough analysis of Reichler et al. (2003, see their Figure 4). Figure 1b shows the same plot using the potential temperature gradient, PTG = dθ/dz < 10 K km^{-1}, to define tropospheric air (Tinney et al., 2022). The PTG tropopause clearly improves the double tropopause identifications compared with LRT, but it still includes some very high O_3 cells near the sub-tropical jet at 30°S.

Figure 1c shows an effective tropopause transition layer (TTL: Hegglin et al., 2009; Homeyer et al., 2010; Pan et al., 2014; Tegtmeier et al., 2020) defined by e90. TTL O_3 values are plotted when model cells containing the e90 = 80 ppb surface occur above the e90 = 90 ppb tropopause cells. From the colors of these TTL cells, we see reasonable tpp O_3 values of 80–160 ppb (Logan, 1999; McPeters et al., 2007) with the large south-to-north gradient expected in July. Age tracers like e90 are able to identify a transition region, but are limited by the model layer thickness.

The O_3 abundance in uppermost tropospheric grid cell (effectively a grid-cell mean) is adopted here as the tpp value. We do not interpolate across stratospheric and tropospheric cells because the sudden change in O_3 vertical gradient as one crosses the tropopause (e.g., Figure 1a) would bring high stratospheric values into the troposphere, dominating the tpp values. The modeled tpp O_3 values on e90 surfaces are not as smooth as might be expected. The high-frequency variability in tpp O_3 shown in Figure 1d, that is, 100 ppb difference across neighboring grid cells, results in part from the sharp tpp gradients combined with the ∼1 km thick CTM grid cells. A major challenge in diagnosing the tropopause region in atmospheric models and satellite observations comes from their relatively coarse vertical resolution. Given the high-frequency variability and sharp gradients seen here in all three dimensions, any tropopause diagnostics cannot usefully be derived from monthly or zonal mean fields. The primary CTM data used here is from 5-day sampling of a 5-year simulation (73 × 5 = 365 global snapshots, averaged into multi-year monthly means to compare with sondes). The potential for large vertical jumps in the tropopause position (Figure 1a) combined with the large vertical gradient in O_3 as one enters the stratosphere

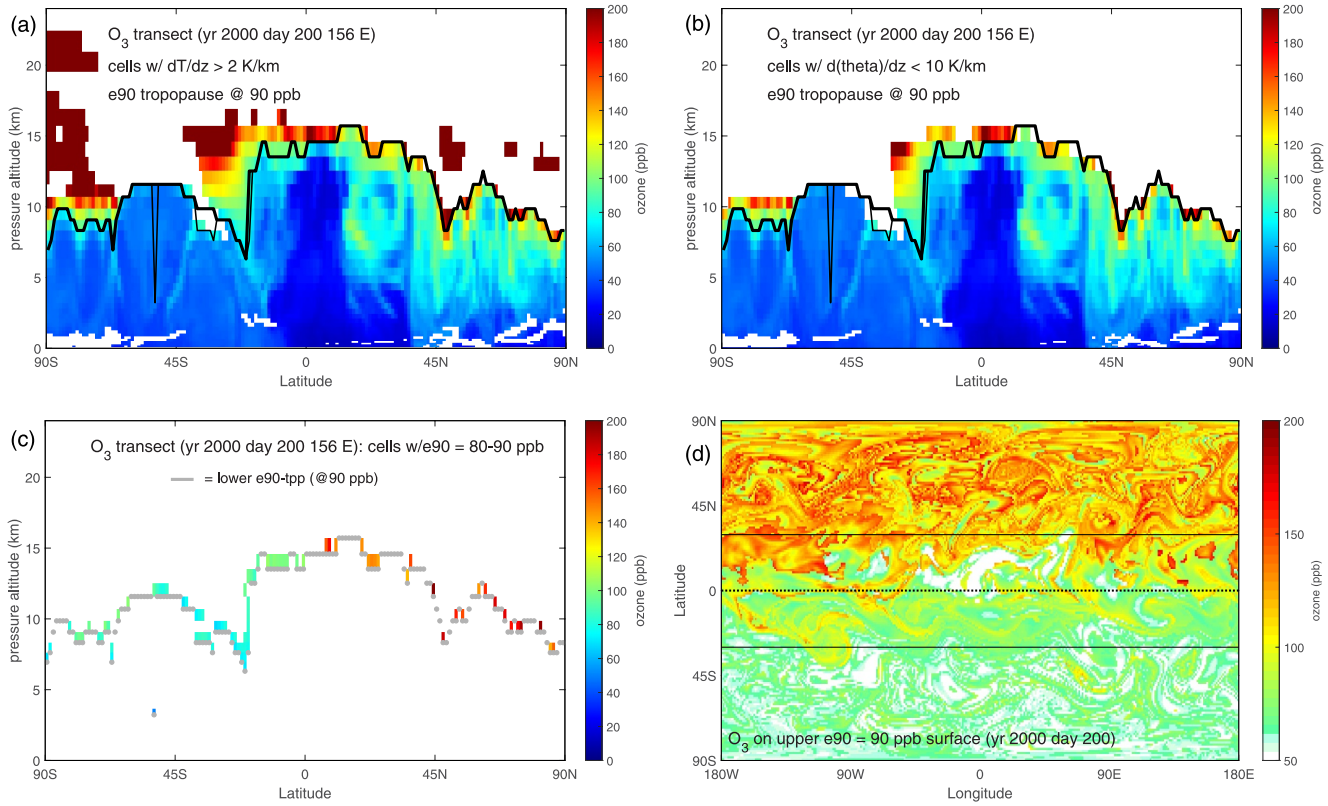

Figure 1. (a) Latitude-by-altitude color map of troposphere-only O_3 (ppb) grid cells from a pole-to-pole transect at longitude 156°E from year 2,000 days 200 (July 18) 00 UTC of a control run of Prather and Zhu (2024), CTM version q76c, IFS cycle 38r1. Non-white grid cells are "tropospheric" based on the World Meteorological Organization (WMO) criteria of the lapse rate being >2 K km^{-1} between that cell and the one above. Low-altitude temperature inversions also show up as white cells. The e90 tropopause is defined at 90 ppb. The upper (thick black line) and lower (thin black line) tropopauses separate (~35°S) where a stratospheric fold occurs. The WMO criteria fails (tropospheric O_3 > 200 ppb) near the sub-tropical jets and winter poles. (b) Latitude-by-altitude color map of troposphere only O_3 for the same transect as in panel (a), except the tropopause criterion is based on the potential temperature gradient, dθ/dz < 10 K km^{-1} (Tinney et al., 2022). (c) Latitude-by-altitude color map of O_3 (ppb) in a hypothetical tropopause transition layer (TTL) for the transect in panel (a). The gray bars correspond to the lower e90 = 90 ppb tropopause in panel (a) (thin black line). TTL O_3 is plotted only for cells where the e90 = 80 ppb tropopause occurs in a cell above. For many latitudes the TTL region is not resolved because of the ~1 km layer thickness. O_3 values are typical of tropopause levels found in the ozonesonde data (70–120 ppb in SH; 130–190 ppb in NH). (d) Longitude-by-latitude color map of tropospheric O_3 (ppb) grid cells on the e90 = 90 ppb surface for the same CTM snapshot as panel (a). Values are not interpolated but taken from the highest tropospheric layer. The high variability, that is, 70–180 ppb in the northern extra-tropics, results in part from the sharp change in vertical O_3 gradient near the tropopause combined with the ~1 km thick CTM grid cells.

means that position errors generate asymmetric O_3 errors favoring very large stratospheric values. Hence, we choose the monthly median value of O_3 as the best statistical tpp values.

2.2. Ozonesondes

Our finest vertical resolution of O_3 at the tropopause comes from ozonesondes, which provide us with a single profile of pressure (p), temperature (T), and O_3 at a resolution of 30–50 m or better (Logan, 1999; McPeters et al., 2007; Thompson et al., 2021). Even with sonde data, the gold standard for locating the tropopause, there are difficulties in defining the tpp location. Reichler et al. (2003) among others have made extensive tropopause comparisons between models and the radiosonde data. We choose the three regions where there they find clear and consistent agreement (northern and southern mid-latitudes plus the tropics) and select three well-established ozonesonde sites from each region: WAL 38°N, ASC 8°S, and LAU 45°S, see Table 1. The Boulder CO ozonesondes are among the best maintained, but the Boulder location is known for tropopause mountain waves affecting the tropopause structure (Siler & Durran, 2015; Worthington, 1998).

With ozonesondes, we have different vertical resolution in reporting that varies over the decades. The coarser resolution of the earliest ozonesondes, 30–50 m (typical of radiosondes), is adequate for this study, and we

Table 1
Observational Data Sets

Type	Source	Files	Profiles	Dates
Ozone	Wallops (WAL, 38°N), SHADOZV06	1,477	1,477/tot	1995–2020
Sondes	Ascension (ASC, 8°S) SHADOZV06	831	831/tot	1998–2022
	Lauder (LAU, 45°S) NIWA	1,973	1,973/tot	1986–2021
OMPS	OMPS-NPP_LP-L2-O3-DAILY_v2.6	3,587	~1,160/day	2014–2023
ACE-FTS	ACEFTS_L2_v4p0_O3.nc (includes O_3, *pp*, T)	1	94,675/tot	2004–2020

process later sondes by averaging 2 to 4 to 6 adjacent points to get a consistent resolution of 30–50 m (see Table 1). From these data, we show in Figure 2 the profiles of the WMO lapse rate (−dT/dz) and PTG (dθ/dz, another tpp criterion proposed by Tinney et al., 2022) for one day from each of our three sites. The small blue dots show these gradients at the ~40 m resolution (adjacent points): the measurement noise in p and T makes it impossible to locate the tropopause using $-dT/dz = 2$ K km^{-1} and $d\theta/dz = 10$ K km^{-1} criteria (thick vertical black lines). If we calculate these gradients using a pair of points separated by about 1 km (red X's, ~25 points apart, skipping the points in between), then the determination of the thresholds is better but still noisy. We then calculate vertical gradients by averaging T, θ, and dz values for 1 km above and 1 km below each point, giving mean values for ½ km above and below, and then differencing these values to get the 1 km gradient at the point (gold O's). We get a smoothly varying curve with a clear and distinct WMO or PTG tropopauses (horizontal black lines), although the two methods do not always agree.

"At the station, the tropopause is determined by visual inspection of the highly resolved temperature profile using the WMO criterion" (Reichler et al., 2003). This approach is not useful nor objective when we need to process thousands of ozonesondes. Thus, we adopt the 1 km above/below averaging to calculate the 1-km gradients at each point. For the WAL sonde #800, the WMO lapse rate criterion clearly identifies a double tropopause (black horizontal lines at 8.3 and 15 km); whereas the dθ/dz criterion finds only the lowermost one. At ASC sonde #600, both methods agree with a single tropopause just above 16 km. The LAU sonde #1100 has a complex profile with three possible WMO tropopauses, and here, the PTG agrees on the lowermost but fails to catch the uppermost. We include the Tinney et al. (2022) criterion specifically because it is based on O_3–H_2O tracer correlations that are clearly a measure of tracer transport barriers. Because our work relies heavily on ozonesondes and vertical gradients and we want a seamless analysis from tropics to mid-latitudes, thus we do not use other dynamical tpp criteria (e.g., potential vorticity (PV) for the mid-latitudes, Turhal et al., 2024).

We calculate multi-year monthly statistics for tpp values at each sonde site. Because of the asymmetry in the O_3 vertical gradient and the occasional large jumps (up or down) in the WMO tropopause as noted for the CTM above, the O_3 tpp values have a very wide distribution skewed to higher values. Rather than trying to remove outliers, which are almost all on the high side, we take the monthly median tpp values.

2.3. Satellite Ozone Profiles

We augment the ozonesonde profiling with satellite measurements, see Table 1. We collected the Ozone Mapping and Profiling Suite (OMPS) Level 2 data for the period 2014–2023 (Kramarova et al., 2024). OMPS reports about 1,200 profiles per day with O_3, T, z* (geopotential height) at regular 1-km levels. The tropopause geopotential height in OMPS is derived from the Global Modeling and Assimilation Office (GMAO) T profiles. The OMPS tpp values for O_3 and T are taken from the uppermost level below the specified tropopause height and, as we do for the CTM, we take the median monthly tpp values over the selected latitude and longitude domains. Because the OMPS data has high-density near-global coverage, it is also used to calculate the tropospheric air mass as a function of latitude over the 60°S–60°N domain using averages for the tpp pressure. Polar night has no OMPS data.

The occultation profiling of the Atmospheric Chemistry Experiment–Fourier Transform Spectrometer (ACE-FTS) provides T, O_3 and p data at regular 1-km intervals (Bernath et al., 2020). We derive the tropopause for each profile from the WMO lapse rate criterion for the first tropopause. The tropopause values for O_3, T, and p are

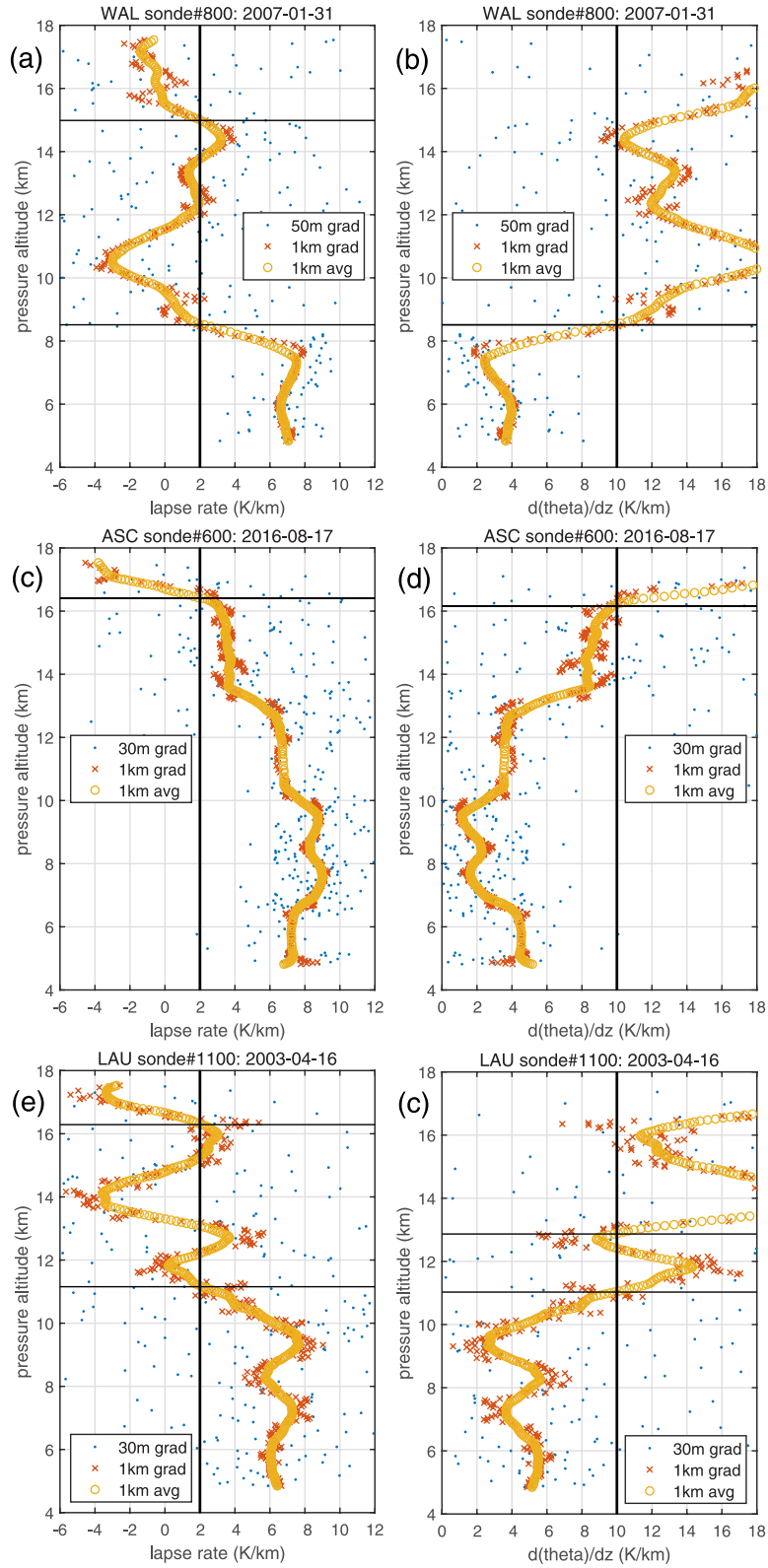

Figure 2.

taken from the uppermost level below the specified tropopause height. With very limited latitude-month coverage for solar occultations in an ACE-FTS orbit (tropical sites have 3-month gaps every year), we use 3-month statistics and all longitudes when sampling for the sonde sites, and the tpp values are simply averaged.

2.4. Matching Models and Measurements

For the CTM data, we define the tropopause in a vertical column as the edge between two adjacent model layers where the lower layer is tropospheric and the upper layer is stratospheric. This tpp edge is chosen using several criteria, see Table 2. The primary tpp criteria is the chemical tracer e90 abundance in each grid cell with thresholds of 70, 80, and 90 ppb. Vertical gradients such as the WMO lapse rate (WMO LR) or PTG are defined using the mean value in each of the layers as the value at mid-layer and the vertical distance (dz) as the separation of the two mid-layers. No attempt is made to curve- or spline-fit T or θ because that can introduce spurious extrema and gradients. These CTM gradient methods are also applied to the satellite profiles. The ozonesonde method uses 1-km averaging described above to calculate the WMO LR or PTG, but then uses the O_3, p, and θ values at the single tropospheric point.

Where multiple tropopauses are found, we calculate the CTM tpp O_3 and p values for both, as shown in later figures. The vertical resolution of the sonde tropopause is about 40 m; that of the UCI CTM is much coarser (0.6–1.0 km at 100–200 hPa); and that of ACE-FTS (∼3 km FOV, Bernath, 2023) and OMPS (1.2 km FOV, Kramarova et al., 2024) are coarser still.

3. Tropospheric Air Mass

3.1. Models

The TAMF can be readily calculated with a global CTM if we can clearly separate tropospheric from stratospheric air masses. The definition of the tropopause or even the transition layer proves ambiguous as noted above. We examine several definitions as described in Table 2. These definitions include both vertical gradient methods (WMO LR, PTG) as well as e90 tracer thresholds, and they include upper and lower tropopauses where there is distinct stratosphere-troposphere folding.

The UCI CTM resolves the tropopause only by layer number in each profile. The tropopause layer numbers from 24 to 33 range from 0.6 to 1.0 km thick, respectively, and cover the pressure range 100–300 hPa. We plot the average layer number of the highest tropospheric layer versus latitude in Figure 3 for the four seasons (DJF, MAM, JJA, SON). The criteria shorthand in the legend are PTG, WMO-LRT, WMO-alt, e/70ppb, e/80ppb, and e/90ppb (see Table 2). For each criteria, the upper tropopauses are shown as O's, and the lower, as X's. In many cases these are indistinguishable. The e90 sequence (green, orange, gold) tracks a smoothly rising tropopause, spanning about 1 layer in the tropics and 2 at the poles. The e/70ppb and e/80ppb show almost no difference between the upper and lower tropopauses, but the e/90ppb (current standard in UCI CTM) shows a layer or more separation in the 45°S–0° for all seasons except MAM. Our interpretation of the formal WMO-LRT (blue) gives the highest tropopause layer with large separation between upper and lower values in the polar regions. Our alternative, simplified WMO-alt (brown) is more similar to the PTG criterion (red) and both are closer to the uppermost e90 criterion (e/70ppb). The latitudinal shape in the tropics shows a clear difference between the gradient (WMO-LRT, WMO-alt, PTG) and tracer methods (e90): the gradient methods have a flat tropopause across 30°S–30°N, but the tracer methods are more rounded. It is not clear what causes this, but the sub-tropical jet is a notoriously difficult region for the gradient criteria (e.g., Reichler et al., 2003, Figure 4).

The annual cycle of TAMF using the six criteria are plotted in Figure 4. We calculated TAMF for the tropics (25°S–25°N), non-polar regions (60°S–60°N), and global. The mean global TAMF values for [WMO-LRT, WMO-alt, PTG, e/70ppb, e/80ppb, e/90ppb] are [0.877, 0.845, 0.852, 0.835, 0.823, 0.808]. The WMO-LRT criterion as implemented for our CTM greatly exaggerates the extent of the troposphere, and the e/90ppb

Figure 2. (a), (c), and (e) Lapse rates (−dT/dz) and (b), (d), and (f) potential temperature gradients (dθ/dz) used to derive the tropopause for sample ozonesondes (dates indicated) from (a) and (b) Wallops VA (WAL), (c) and (d) Ascension I. (ASC) and (e) and (f) Lauder NZ (LAU). The gradients calculated from neighboring pairs of points (30–50 m apart) are shown as blue dots; those calculated from pairs 1 km apart, as red x's; and those calculated from averaging T and θ over a km above and below each point, generating a 1-km gradient at the point, as gold circles. Troposphere values are defined as LR > 2 and dθ/dz < 10 K km^{-1}. Upper and lower tropopauses are shown as horizontal lines, see Table 2.

Table 2
CTM and the Different Tropopause Definitions

CTM grid cells	The chemistry-transport model here has a regular latitude-by-longitude 160 × 320 grid of resolution ∼1.1° and 57 eta-coordinate levels defining the pressure edges of each layer based on the T159L60N80 grid of the ECMWF IFS model. The lowest 5 IFS L60 layers are combined into 2 CTM near-surface layers. The vertical resolution of model layers in the upper-troposphere/lower-stratosphere increases regularly from 0.8 km near 10 km altitude to 1.1 at 17 km altitude. The CTM has layer numbers increasing upward.
e90 tracer	Synthetic chemical tracer, emitted uniformly everywhere from the surface, decaying with a 90-day e-fold time. Calculated in the CTM for every cell and every time step. Emissions are set to give an atmospheric mean abundance of 100 ppb. This techniques has been used to study O3 seasonality at the tpp (Prather et al., 2011)
• e/90ppbU	Highest altitude model layer with e90 > 90 ppb is designated the uppermost tropospheric layer, and its upper boundary is the upper tropopause. The tropopause value for O3 and *TT* is the mean value of the uppermost tropospheric layer. No interpolation is done because of the large change in dO_3/dz.
• e/90ppbL	Lowest altitude model layer with e90 ≤ 90 ppb is designated the lowest stratospheric layer, and its lower boundary is the lower tropopause. Often the same as e/90ppbU.
• e/80ppbU	Highest altitude model layer with e90 > 80 ppb, as above
• e/80ppbL	Lowest altitude model layer with e90 ≤ 80 ppb, as above
• e/70ppbU	Highest altitude model layer with e90 > 70 ppb, as above
• e/70ppbL	Lowest altitude model layer with e90 ≤ 70 ppb, as above
Lapse Rate Tropopause (LRT)	Lapse Rate (LR = -dT/dz, K/km) is calculated between 2 vertically aligned layers using the mean temperatures of each layer and the mid-layer altitude. The LR values and thresholds (LR < 2 K/km = strat) apply to the upper boundary of the lower box used to calculate it.
• WMO-LRTU	WMO LR Tropopause (Upper). Here we try to match the WMO definition as closely while following the logic of the PTG algorithm. The Lapse Rate (LR = -dT/dz, K/km) is calculated between 2 adjacent layers using the mean temperatures of each layer and the mid-layer altitude. Two LRs for layer *kk* are calculated: LR1 = -[T (*kk*+1)−T(*k*)]/[z (*kk*+1)−z(*k*)]; and LR2 = -[T (*kk*+2)−T(*k*)]/[z (*kk*+2)−z(*k*)]. Thus, LR1 spans ∼1 km above the mid-point of layer *kk* and LR2 spans ∼2 km above. All layers below 4.4 km are set as tropospheric (trop). All layers above 31 km are stratospheric (strat). Define strat layers as meeting the criteria LR1 <2 K/km. If the first (lowermost) strat layer has 2 trop layers immediately above, then relabel it trop and keep going. Having found the first strat layer that passes this test, we check that either LR2 from the current layer or LR1 from the layer above meets the <2 K/km criterion. If these conditions are met, we have the first strat layer and the tpp is the lower boundary of that layer. Now look for a 2nd tpp above: If there is an LR1 ≥ 2 K/km above (i.e., trop layer) then define the 2nd tpp (WMO-LRTU) as the lower boundary of the next strat layer. Otherwise, the 2nd tpp is the same as the 1st.
• WMO-LRTL	WMO LR Tropopause (Lower). This is just the 1st tpp from WMO-LRTU.
• WMO-altU	Alternate WMO Upper tpp. Similar to WMO-LRT but easier to implement with CTM grid. Identify trop layers from LR > 2 K/km. Set layers below 6 km altitude to trop. Set layers with θ > 500 K to strat. From the bottom, find the first trop layer with a gap of 2 strat layers (∼2 km thick region) above it. This is the 2nd (Upper) tpp. Repeat looking for a trop layer with a gap of at least 1 strat layer above. This is the 1st (Lower) tpp. Both tpp can, and often are, the same, but strat-trop folds resolved by the CTM are easily identified.
• WMO-altL	Alternate WMO Lower tpp. The 1st (Lower) tpp from WMO-altU search above.
Potential Temperature Gradient Tropopause (PTG)	The PTG (dθ/dz, K/km) is calculated like the LR between 2 vertically aligned layers using the mean θ of each layer and the mid-layer altitude. The PTG values and thresholds (PTG ≥10 K/km) apply to the upper boundary of the lower box used to calculate it.
• PTG-tppU	PTG Tropopause (Upper). The algorithm here is similar to the original paper (Tinney et al., 2022), adapted to work efficiently with the CTM diagnostics. The PTG (dθ/dz, K/km) is calculated between 2 adjacent layers using the mean θ of each layer and the mid-layer altitude. Two PTGs for layer *kk* are calculated: PTG1 = [θ(*kk*+1)−θ(*k*)]/[z (*kk*+1)−z

Table 2
Continued

	(k)]; and PTG2 = [θ(kk+2)−θ(k)]/[z (kk+2)−z(k)]. Thus, PTG1 spans ∼1 km above the mid-point of layer kk and PTG2 spans ∼2 km above. All layers below 4.4 km are set as trop. All layers above 31 km are strat.
	Define strat layers as meeting the criteria PTG ≥10 K/km. If the first (lowermost) strat layer has 2 trop layers immediately above, then relabel it trop and keep going. Having found the first strat layer that passes this test, we check that either PTG2 from the current layer or PTG1 from the layer above meets the ≥10 K/km criterion. If these conditions are met, we have the first strat layer and the 1st tpp is the lower boundary of that layer.
	Now look for a 2nd tpp: If we find a trop layer above the strat layer above the 1st tpp, then we search for the first strat layer above it using a revised threshold, PTG1 ≥ 15 K/km. This identifies the 2nd tpp (PTG-tppU) as the lower boundary of the next strat layer. Otherwise, the 2nd tpp is the same as the 1st.
• PTG-tppL	PTG Tropopause (Lower). This is just the 1st tpp from the PTG-tppU search above.

current standard has too small a troposphere. This result is supported by the tpp P at the three sonde sites shown later where the WMO-alt is easy to implement in our CTM and matches the OMPS tpp P (effectively the GMAO model's WMO tropopause). In the tropics OMPS values lie between the e/70ppb and e/80ppb values; while in the non-polar regions they align with the e/80ppb values. If we accept the OMPS TAMF values, then the global TAMF lies in the range of 82%–83% (orange and gold points). We do not understand why our WMO LR methods applied to the UCI CTM give such different results from the OMPS results, also presumably using the WMO

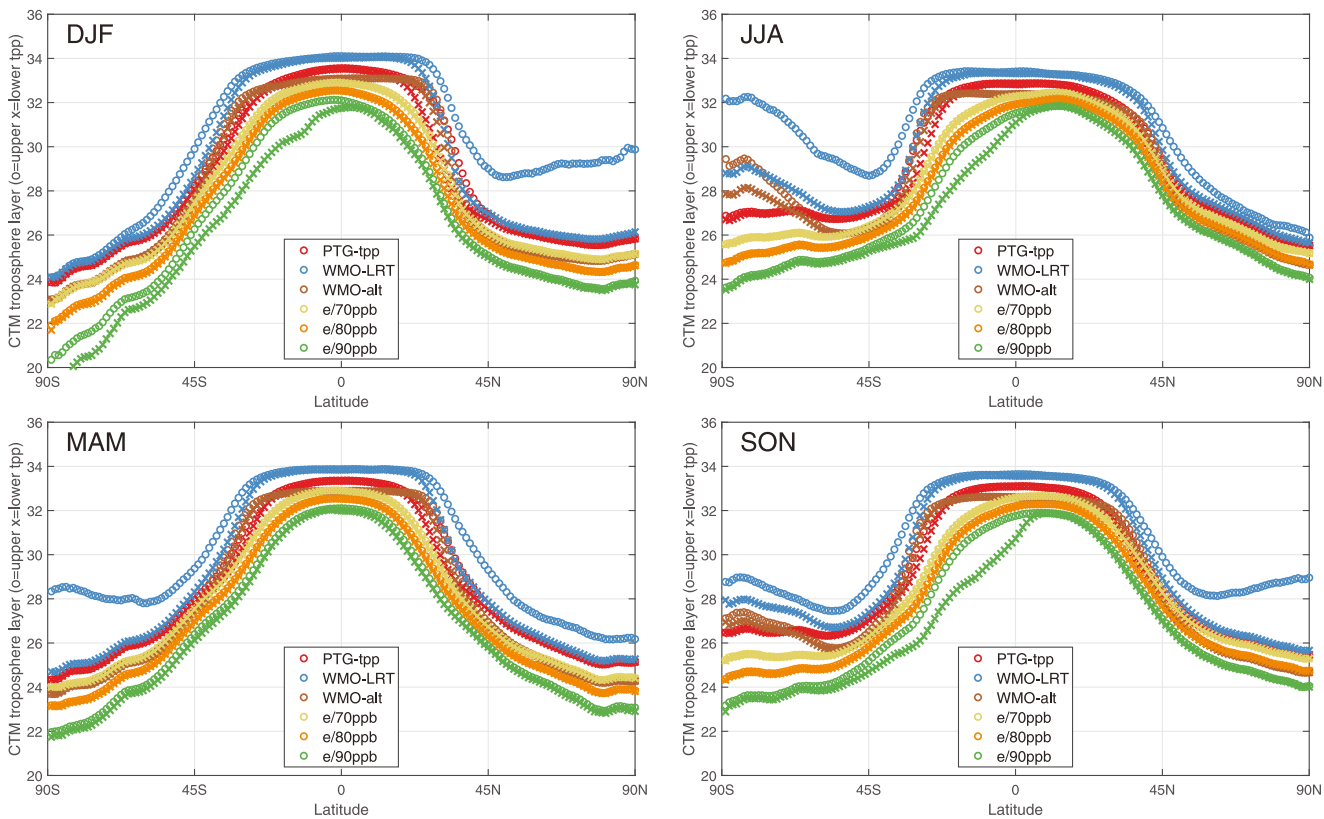

Figure 3. Mean CTM tropopause layer number versus latitude for the four seasons: DJF, MAM, JJA, SON. The layer refers to the grid cell just below the tropopause and is calculated using several criteria: (blue) World Meteorological Organization lapse rate tropopause; (red) potential temperature gradient, PTG; (brown) WMO alternate LR method; (green) e90 tracer = 70 ppb; (orange) e90 = 80 ppb; and (gold) e90 = 90 ppb (gold). Typically, the top of layer 24 is 302 hPa; that of layer 26 is 242 hPa; and that of layer 32 is 104 hPa. The lower and upper tropopause (x's and o's respectively) are clearly separated except for the e90 = 70 and 80 ppb thresholds, which for the most part find only single tropopauses. These CTM values are derived from 5 years (2000–2004) of 5-day snapshots for each season (approximately 90 global samples). See text.

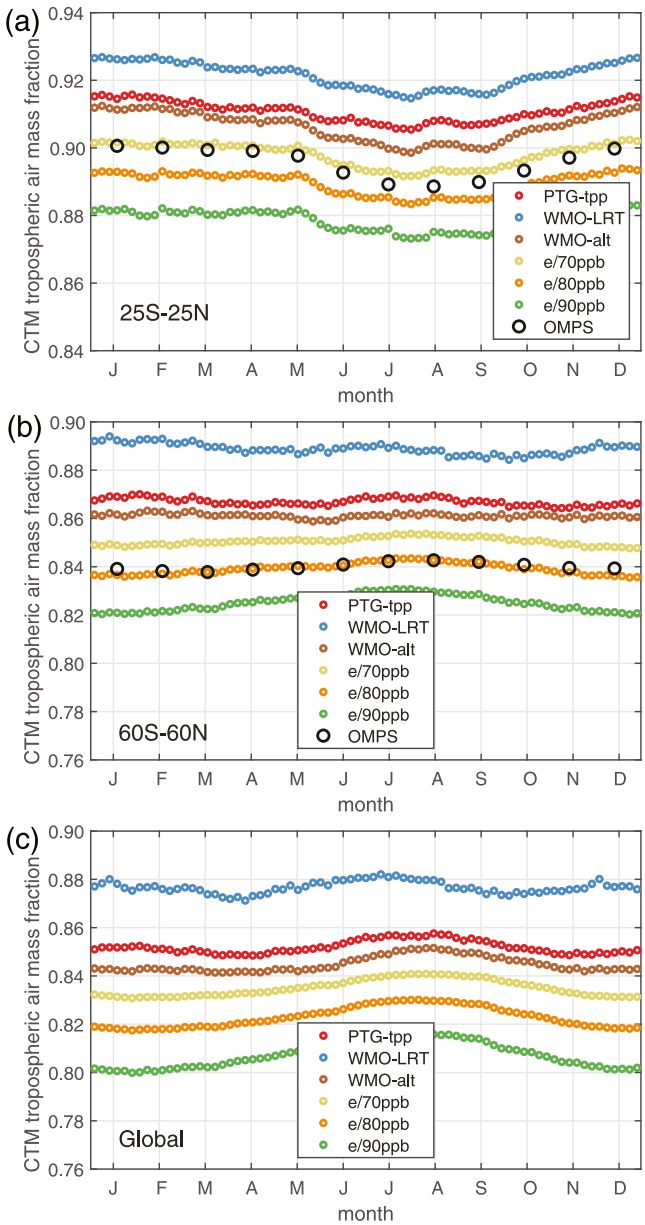

Figure 4. Tropospheric air mass fraction for (a) 25°S–25°N, (b) 60°S–60°N, and (c) globally throughout the year using six different criteria for the tropopause, see text. Results are based on 5-day sampling of the CTM over the years 2000–2004. Also shown are the monthly fractions from the Ozone Mapping and Profiling Suite satellite data (large black O's) for 2014–2023, which does not have global coverage.

method, because both meteorological data sets are based on assimilated/forecast fields with similar vertical resolution.

3.2. Sondes

Our three ozonesondes sites, covering distinct regions, provide a long-term record of tpp pressure and O_3. Given the individuality of each sonde site and each sonde profile, we combed through each of these, removing tagged or obviously bad data. The sequence of processing is given below:

1. *Find O_3 NaNs:* Flagged (e.g., −999); negative; too large (>20 ppm); pressure (<1 hPa)
2. *Find T NaNs:* Flagged; $T < 160$ K
3. *Collapse profile*: Remove: O3 NaN pts; yoyo (down-up) sections; descent part of profile; all $p < 10$ hPa.
4. *Drop whole profiles:* With <25 O3 points; with top $p > 65$ hPa = ASC (44), LAU (24), WAL(58).
5. *Reduce # pts:* Many profiles have 6,000–8,000 pts and resolution <10 m; create 4-pt (ASC and WAL) and 6-pt (LAU) averages to get 30–50 m spacing for p = 70–500 hPa.
6. *Calculate* 100-m *gradients:* Use centered 2nd-order finite difference to get 100-m values for LR = −dT/dz and PTG = dθ/dz at each point. These are too noisy to identify tropopause, see Figure 2
7. *Calculate stable* 1-km *gradients:* Average z, T, O3 for 20 points (~1 km) below/above each point; difference these to get a smoothed 1-km-average LR and PTG, see Figure 2. Due to the smoothed 1-km resolution profile, we do not apply the WMO criterion that gradients must be sustained above the threshold point.
8. *Lower tropopause:* LRT = point below lowest LR ≤ 2 K/km, limited to $500 > p > 80$ hPa; PTG = point below lowest dθ/dz ≥ 10 K/km, ibid
9. *Upper tropopause:* LRT = highest LR > 2 K/km, limited to $500 > p > 80$ hPa; PTG = highest dθ/dz < 10 K/km, ibid;

Thresholds are not changed for the upper PTG as in Tinney et al., 2022. Several studies of radiosonde data have shown that the tropopause is rising at rates of 1–2 hPa per decade over the periods used in this study (Santer et al., 2003; Seidel & Randel, 2006; Xian & Homeyer, 2019). The differences across models, measurements, and metrics are much larger, >10 hPa; and hence we did not try to remove any trends from the models or measurements.

Comparisons of sonde data with CTM, OMPS, or ACE use a geographic block and monthly averages over the multiple years. CTM values are calculated for grid-cell blocks centered on each ozonesonde site: 32°–44°-N × 240°−360°E for WAL; 14°–2°S × 285°−45°E for ASC; 51°–39°-S × 111°−230°E for LAU. Using 5 years (2000–2004) of 5-day snapshots, we have approximately 30 global samples for each month. CTM results use median values for tpp p and O_3 in each block and then average over the 30 blocks. The OMPS profiles are dense in latitude and the sonde geographic blocks are chosen with the same 12° latitude range about the sonde site but with a width of only 60° in longitude, producing a total of 90,138, 87,892, and 87,081 total profiles (all months) for WAL, ASC, and LAU, respectively. For OMPS tropopause values, we also collect only the median values of the latitude-longitude block. ACE-FTS data are sparse, and so we adopt the same latitude ranges but accept all longitudes and average over 3 months. For each site and season the number of ACE-FTS profiles varies from 265 to 994, and because of this sparse coverage, we used mean rather than median values for tpp P and O_3. This choice may affect tpp O_3 more than P as seen below.

The tpp pressure is a direct measure of the TAMF. Figure 5a shows the monthly median tpp pressure for the three sonde sites. The OMPS (X's) and ACE-FTS (boxes) agree well with the sonde values (O's), where sonde and

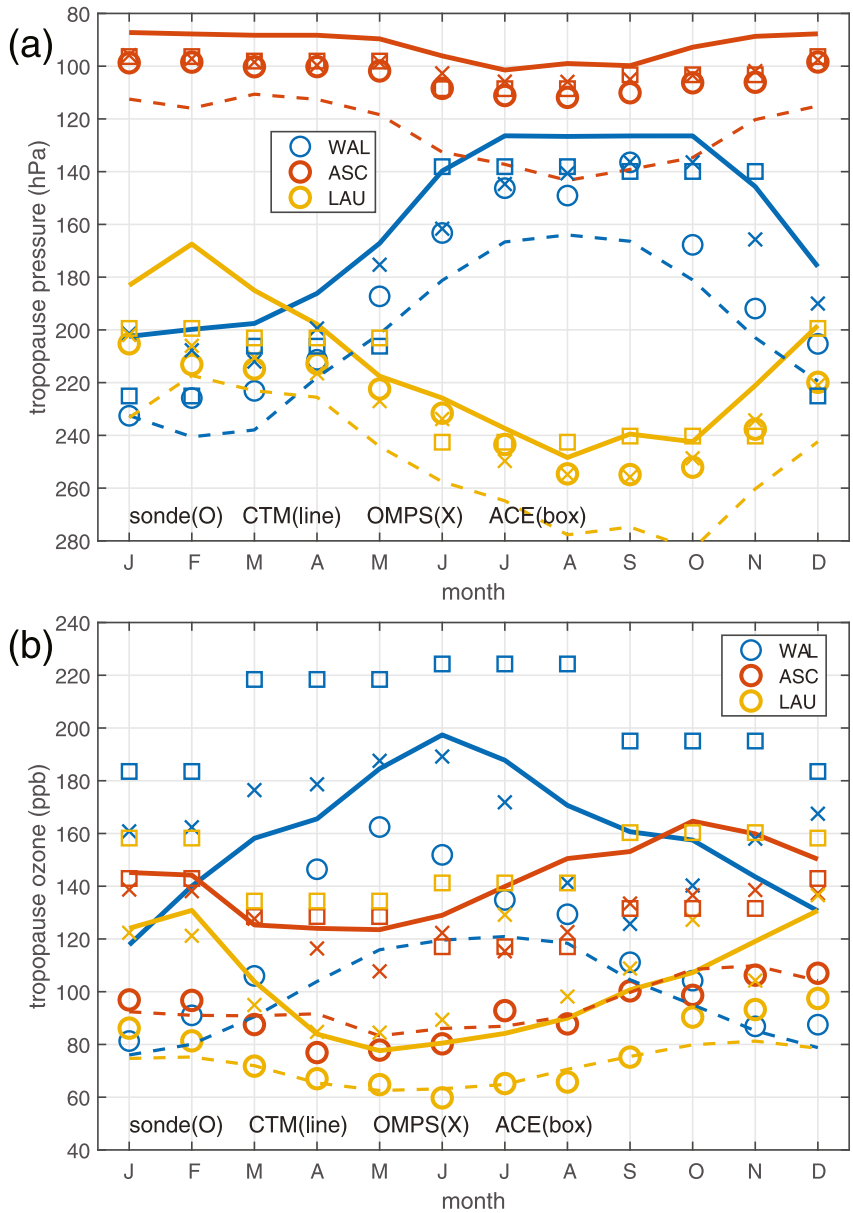

Figure 5. (a) Monthly tpp pressure (hPa) at three ozonesonde sites (WAL, ASC, LAU) derived from ozonesondes (o), CTM (lines), Ozone Mapping and Profiling Suite (x), and Atmospheric Chemistry Experiment - Fourier Transform Spectrometer (box). The tpp definition for sondes and satellites uses the lower lapse rate tropopause. For CTM the solid line uses lower World Meteorological Organization-alt criterion, and the dashed line uses e90 = 90 ppb criterion. See text for details. (b) Monthly tpp ozone (ppb), as for pressure.

satellite use the lower WMO LRT given the difficulties with the upper LRT (Figure 1a). The CTM range shows two tpp criteria (solid line, WMO-alt; dashed line, e/90ppb), which fall distinctly above and below the sonde values.

A deeper look at the sonde tpp pressure versus the six CTM criteria is shown in Figures 6a–6c. The sonde pressures fall mostly in parallel with the e90 tracer surfaces and supports the inference from Figure 4, wherein the best match with the OMPS TAMF would be a criteria somewhere between e/70ppb (orange) and e/80ppb (gold). The gradient methods as implemented for the CTM are unable to match the sonde tropopause.

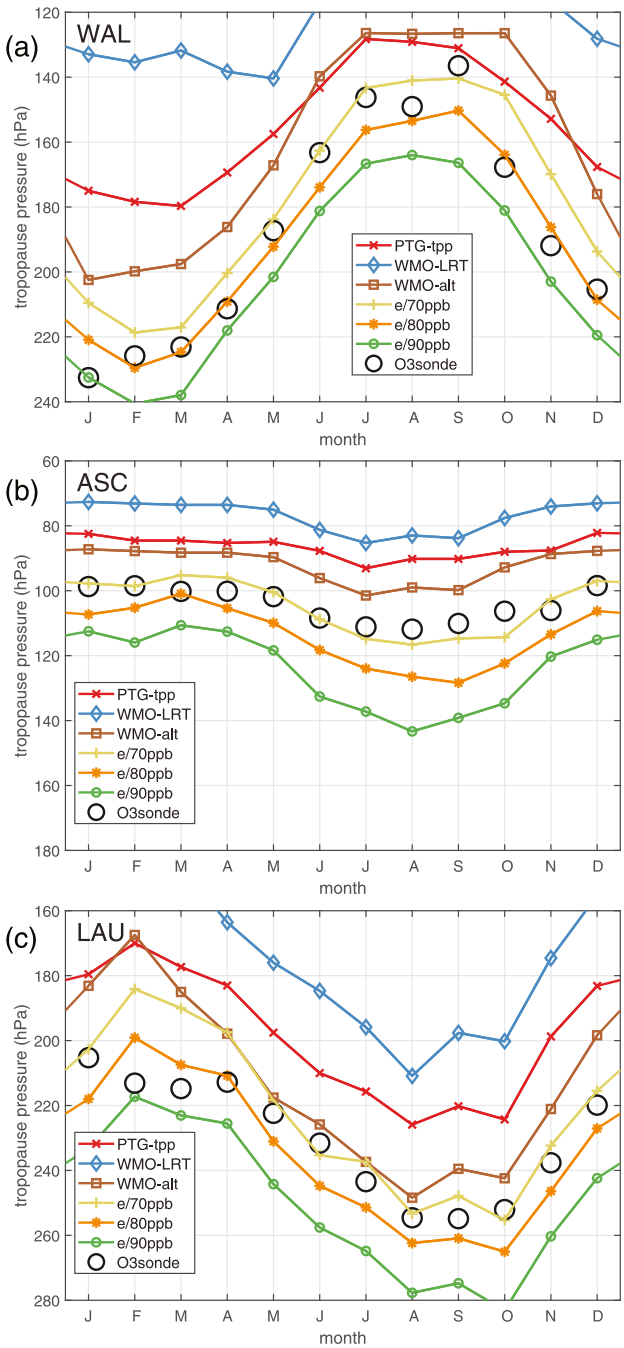

Figure 6. (a–c) Comparison of monthly median lower World Meteorological Organization lapse rate tropopause tpp pressure (hPa) from sondes (black O's) versus the 6 CTM criteria (colored lines with symbols, see legend). Results are shown for (a) Wallops (WAL), (b) Ascension (ASC), and (c) Lauder LAU). See text.

4. Tropopause O_3

The tpp O_3 (Figure 5b) shows a more chaotic picture than its partner figure of tpp pressure. What is unusual with the satellite-sonde comparison is that all three agree on tpp P but have widely disparate tpp O_3. The consistency in this figure is that a large high bias in tpp O_3 values is found for both satellites at all three sites. For OMPS (X's), the [WAL (blue), ASC (red), LAU (gold)] bias is [+47 ± 26, +35 ± 6, 32 ± 14] ppb, where ± is the standard deviation over the 12 months. For ACE-FTS (boxes), the bias is higher still, [+89 ± 16, +37 ± 9, 72 ± 7] ppb. All three measurements are well calibrated and this cannot be a problem with locating the tropopause, and so it must be a problem with resolving the O_3 profile. We analyze the bias errors associated with vertical averaging of the satellite instruments. Calculating the average O_3 that lies ±1 and ±2 km above and below the tropopause, we find as expected that the sharp change in dO_3/dz about the tropopause produces a high bias for [WAL, ASC, LAU] of [+18, +1, +19] ppb for 1 km and [+43, +7, +39] ppb for 2 km. The lack of high bias with vertical averaging at ASC can be explained by the more extensive tropical transition layer and the much smaller change in dO_3/dz. So, the satellite biases at WAL and LAU can be understood from the vertical resolution, and the still larger ACE-FTS differences are probably due to the use of mean tpp O_3 instead of median. As an informative check on this, we calculated ACE-FTS median tpp O_3 (as for the others) and find a much-reduced discrepancy of [+65 ± 17, +19 ± 7, 55 ± 8] ppb, more inline with OMPS and consistent with a 2-km field of view. Thus, it is critical to use median tpp O_3 values to avoid the large asymmetric errors induced by tpp P errors.

The CTM tpp O_3 may not match the sondes due to any number of flaws in tropospheric and stratospheric chemistry modeling, but we do not expect the vertical averaging problem found with the satellites because we are not averaging or interpolating O_3 in the vertical but selecting the O_3 value from the highest fully tropospheric model layer. Because the UCI CTM uses 3-D s-order moments for tracer transport, including in the vertical, there is little numerical mixing across the tropopause (Lauritzen et al., 2014; Prather, 1986). CTM e/90ppb (dashed lines in Figure 5b) comes close to match the sonde values (O's), but the CTM WMO-alt, which better matched the tpp pressures, has higher tpp O_3 values, more like OMPS (X's).

A deeper look at the sonde tpp O_3 versus the six CTM criteria is shown in Figures 7a–7c. The e/90ppb criteria matches much of the sonde data, but for WAL and summertime LAU, where the better criteria would be e/80ppb or even e/70ppb. The light gray O's show the impact of ±1 km averaging of the sonde data discussed above and indicate that this effect is similar to a 10ppb lower shift in the e90 criteria.

All of the gradient criteria applied to the CTM clearly overestimate tpp O_3. This flaw is even worse when one considers regions other than our carefully selected sonde sites. Tropopause O_3 for all latitudes and four seasons is shown Figure 8. Upper (O) and lower (X) tropopauses for each of the six criteria are plotted, but only distinguishable for the gradient methods. The three e90 criteria show a smooth latitude-by-season pattern that is continuous across latitudes with regularly increasing O_3 values from e/90ppb (green) to e/80ppb (orange) to e/70 ppb (purple). All e90 thresholds show a clear latitude-by-season pattern that evolves as expected from the sonde data (Logan, 1999; McPeters et al., 2007). It is interesting is that these e90 isopleths separate more during the summer season and contract during winter, which indicates a different rate of vertical mixing in these seasons. The gradient methods (PTG, WMO-LRT, WMO-alt) simply blow up at the sub-tropical jets and the poles. One of the gradient methods (WMO-alt) is almost consistent with

10.1029/2025AV001651

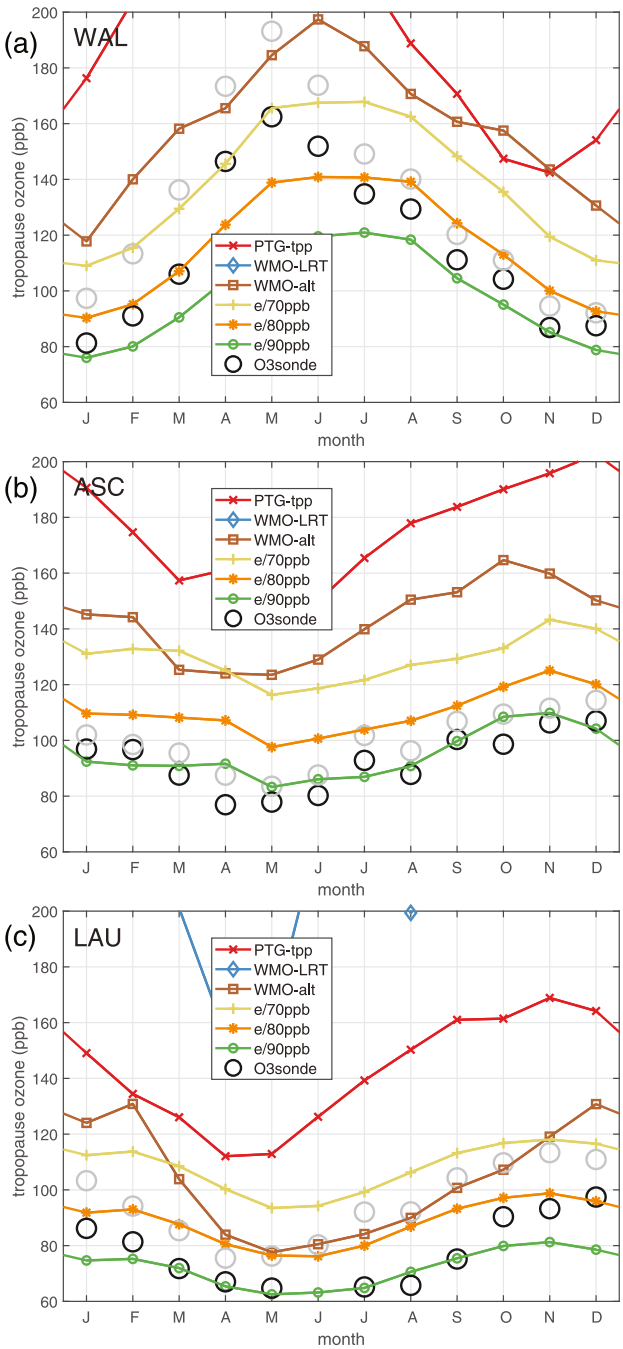

Figure 7. (a–c) Comparison of monthly median lower World Meteorological Organization (WMO) lapse rate tropopause (LRT) tpp O_3 (ppb) from sondes (black O's) versus the 6 CTM criteria (colored lines with symbols, see legend). Results are shown for (a) Wallops (WAL), (b) Ascension (ASC), and (c) Lauder LAU). See text. The gray O's show sonde O_3 averaged ±1 km about the tropopause. For WAL and ASC, the WMO-LRT tpp O_3 values (red, not seen) are off the top of the graph.

the e90 methods for latitudes near our three selected sites. We accept that our CTM implementation of WMO-LRT and PTG may not be optimal, but the total blow-up of these methods near the jets precludes their use in any statistical analysis of global tpp O_3, even with median O_3 as used here. This problem was evident in sample transect of Figure 1.

5. Calibrating the e90 Tropopause

How does one calibrate the e90 tpp threshold? The obvious choice is to select the tpp P (Figure 6) which gives a fairly straightforward threshold of 70–80 ppb based on the three sonde sites. This leads to a too-large tpp O_3 (Figure 7), but we can readily accept that the modeled O_3 has biases of 20–30 ppb. There is, however, another possibility. Suppose the CTM tracer transport (large-scale plus convection plus diffusion plus numerical mixing) is too weak and the transition from mixed to stably stratified occurs at a smaller TAMF such as at e90 = 90 ppb.

We find this effect when comparing two successive versions of our baseline meteorological data: IFS Cycle 38r1 and Cycle 43r3, see Table 3. The two cycles have differently sized tropospheres for the same e90 threshold: the TAMF increases about 0.7% (absolute units) from Cycle 38r1 to Cycle 43r3. The stratospheric circulation also changes, yielding more stratospheric (+9 DU) and tropospheric O_3 (+3 DU, numbers refer to 70–80 ppb thresholds). Lightning NOx generation increases with Cycle 43r3, indicating more/deeper convection, also consistent with an expanded troposphere. Other interesting results in Table 3: stratospheric folds are more readily discerned for the 90 ppb threshold, making up 0.3% of the total air mass, than for 70–80 ppb (0.05%). In the transition zone 70–90 ppb, there is about 3 DU of O_3, about 10% of the total tropospheric burden.

If we calibrate e90 using the Cycle 38r1 model against the sonde tpp pressure, we get about a 75 ppb threshold and a TAMF of 83.0%. But that e90 threshold for Cycle 43r3 would give us a TAMF of 83.7%. When considering the diversity of CTMs and meteorological data, one can expect even larger differences. We need a way to calibrate against mixing barriers that allows each model version to generate its own TAMF, rather than using sonde tpp P to force the tropopause definition.

The e90 tracer provides us with a quantitative measure of transport time, for example, the difference between the 81 and 79 ppb isopleths is 90 ln (81/79) = 2.2 days. The vector spatial gradient of ln (e90), ∇·ln (e90), can be largest in any direction (height, latitude, longitude), but here we assume it is predominantly vertical, d (ln (e90))/d(z). Our assumption is valid only in certain regimes (core tropics and mid-latitudes) and fails when the mixing barriers are horizontal as at 18°S in Figure 1a. An increase in d (ln (e90))/d(z) indicates a transport barrier, and by using the decay time we can calculate the e90 mixing gradient G = 90 d ays(ln (e90))/d(z) in day km^{-1}. The inverse, 1/G, gives a mixing velocity, effectively the rate at which the e90 surface must be propagating into the stratosphere in km per day.

The square of the Brunt-Väisälä frequency ($N^2 \equiv$ g d (ln(θ))/d(z), s^{-2}) is a measure of stability against vertical displacement. We define a new index combining N^2 and 1/G, d (ln(θ))/d (ln (e90)) (dimensionless). It is possible to derive this index from sondes or aircraft if we measured a species that gives age-of-air. A plot of the annual model mean of d (ln(θ))/d (ln (e90)) versus e90 at the three sonde sites is shown in Figure 9a (dashed lines). We place a horizontal gray line at e90 = 81 ppb where d (ln(θ))/d (ln (e90)) stabilizes below 0.15. This threshold identifies a clear transition in atmospheric mixing, but what transition? The tropopause has also been defined as where N^2

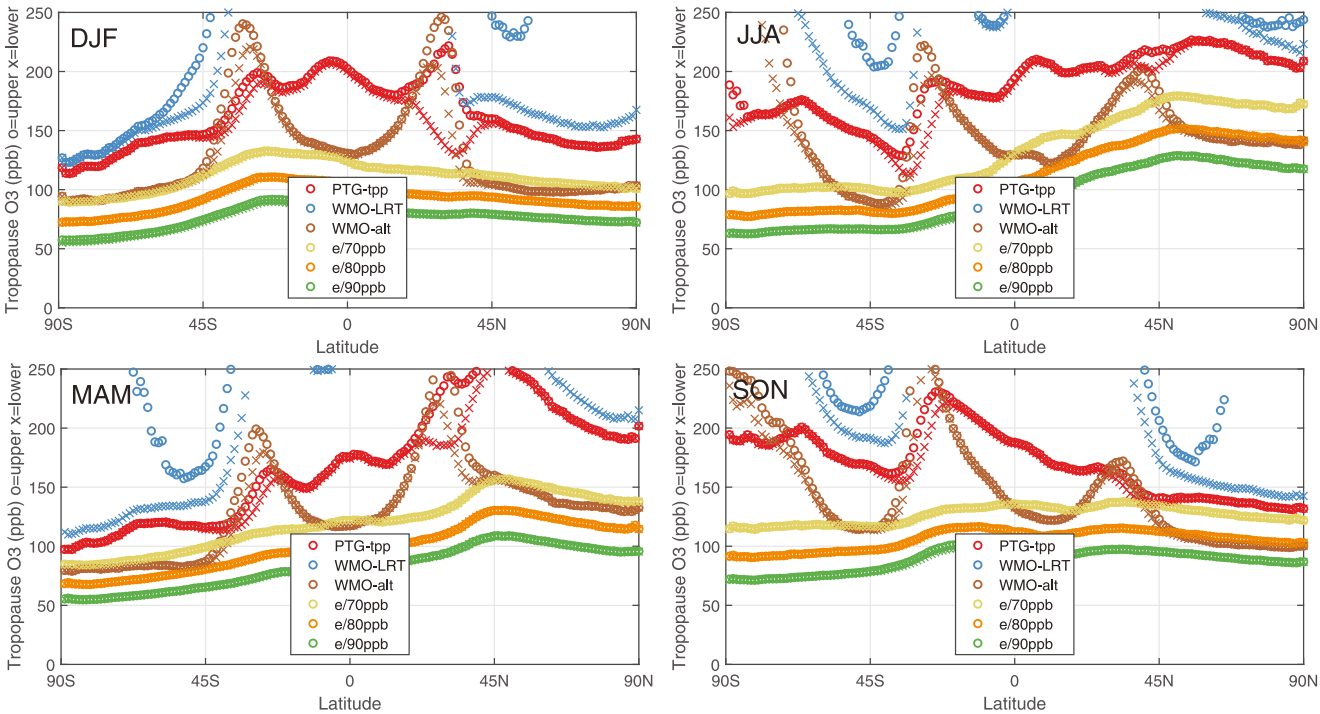

Figure 8. Tropopause O_3 (ppb) versus latitude for four seasons (DJF, MAM, JJA, SON) using the 6 CTM tropopause criteria. Results are derived from median O_3 values from the 5-day sampling over 5 years. Upper (o) and lower (x) tropopauses are shown but are distinguishable only for the potential temperature gradient and World Meteorological Organization criteria.

jumps from about 1 to >3 × 10^{-4} s^{-2} (Homeyer et al., 2010; Tinney et al., 2022). This threshold is at $N = 0.017$ s^{-1} and is shown by the vertical gray line in Figure 9a. Values of N versus e90 are also plotted (O's). Both measures point to the same e90 tpp value for the three sites.

The e90 atmospheric structure is shown in Figure 9b using the same e90 y-axis. The e90 altitude (multi-year annual mean) is plotted as dashed lines for each sonde site. Using e90 = 81 ppb (horizontal gray line), the tpp altitudes for LAU, WAL, ASC are 10, 12, and 16 km, respectively. The vertical e90 age-of-air gradient (days per km) are plotted (X's), and at the e90 = 81 ppb line we see a clear transition to slow, stratospheric vertical transport times of 16–30 days per km.

From this analysis, we find that a consistent e90 threshold, about 80 ± 5 ppb, roughly follows the LRT across the three latitudes of our sonde sites and throughout the year. Further, a threshold of 81 ppb identifies a distinct mixing barrier at the three sites. Can we identify a similar universal tpp threshold for other non-synthetic (measurable) atmospheric tracers? Defining the tropopause boundary using dynamical and chemical tracers such as PV (K m^2 kg^{-1} s^{-1}) and O_3 (ppb) is well established (e.g., Bethan et al., 1996; Birner, 2006; Schoeberl, 2004; Tinney et al., 2022; Wirth, 2000). Using our CTM with our newest set of meteorological data (Our ECMWF IFS cycle 43r3 forecasts saved the PV at each cell), we sought a single e90-PV or e90-O_3 relationship for the tropopause region. Unfortunately, we found these relationships varied with latitude and season and did not provide a universal tpp threshold. This result is not new: Schoeberl (2004) notes that "unfortunately, no standard definition for the extratropical PV tropopause exists"; and Logan (1999) showed the large seasonal-latitudinal variations in tpp O_3 found here in Figure 7. This failed effort is documented in the Supporting Information S1.

6. Recommendations

We cannot and should not expect different models to have the same sized troposphere (TAMF), so a best approach might be to have models calibrate the e90 threshold for their own tracer-transport modules. The vertical transport gradient G in the UCI CTM shows a clear and reasonably sharp transition at 15 days per km for the three sonde sites, and thus identifies the e90 tpp transition at 81 ± 2 ppb with a TAMF of 82.2 ± 0.3% for IFS Cycle 38r1. This

Table 3
Ozone and Air Mass Diagnostics From CTM Simulations

met data (ECMWF)	IFS cycle 38r1 T159L60N80 grid	IFS cycle 43r3 T159L60N80 grid
CTM model	T159L57N80: IFS lowest layers 1:3 and 4:5 combined into CTM layers 1 and 2	
e90 @ 90ppb		
O3 total (DU)	309.3	321.3
O3 strat (DU)	277.6	287.4
O3 trop (DU)	31.7	33.9
O3 folds (%/trop)	0.6%	0.5%
trop dry air mass	80.8%	81.3%
folds dry air mass	0.29%	0.19%
e90 @ 80ppb		
O3 total (DU)	309.3	321.3
O3 strat (DU)	276.3	285.6
O3 trop (DU)	33.0	35.7
O3 folds (%/trop)	0.2%	0.3%
trop dry air mass	82.3%	83.0%
folds dry air mass	0.05%	0.07%
e90 @ 70ppb		
O3 total (DU)	309.3	321.3
O3 strat (DU)	274.9	283.7
O3 trop (DU)	34.4	37.6
O3 folds (%/trop)	0.2%	0.3%
trop air mass	83.6%	84.3%
folds air mass	0.04%	0.06%

Note. Values calculated from simple 36-month average of first-of-month snapshots of tracers and dry air mass from 1 February 2001 through 1 January 2004, and may differ slightly from similar statistics derived from the 5-day snapshots. Tropospheric (trop) O_3 and air mass are calculated from the uppermost trop layer and include folds. All runs used CTM version q7.7 (23 August 2024) and the threshold for switching between tropospheric and stratospheric chemistry was kept at 90 pbb, while the post-run diagnostics used the 3 criteria noted above.

calibration method will not work near the sub-tropical or polar jets because the major transport barriers are horizontal, across the jets. That it works in the core tropics and mid-latitudes, providing similar e90 thresholds, is encouraging. There is no need to restrict the model calibration to specific sonde sites as done here, but any regions away from the jets where the vertical e90 gradient measure the mixing rate will do.

The sonde tpp P values then provide a verification data set. We compare the modeled tpp P using the e90 calibration with the sonde WMO LRT tpp P. This comparison should be done with the vastly greater radiosonde network of 340 sites as in Reichler et al. (2003). Their work also shows that these comparisons work best in locations where the mixing barrier is primarily vertical, see the marked off jet regions in their Figure 4. How well does our CTM compare for the three sites? From Figure 6, we see that our e90 = 81 ppb tropopause matches most of the months at WAL (except June, July, August, and September), but for these months and most months at ASC and LAU, the sonde tpp P is much closer to e90 = 70 ppb values. We surmise that our Cycle 38r1 tracer-transport model lacks deep convective mixing. This analysis needs to be reexamined with the Reichler data.

What did we learn about tpp O_3? The best comparisons between observation and model, and between different observations, require Level 2 data (single profiles) and snapshot model sampling. The uncertainty or error in determining the tpp location results in a large high bias for the mean tpp O_3, and thus medians or other statistical

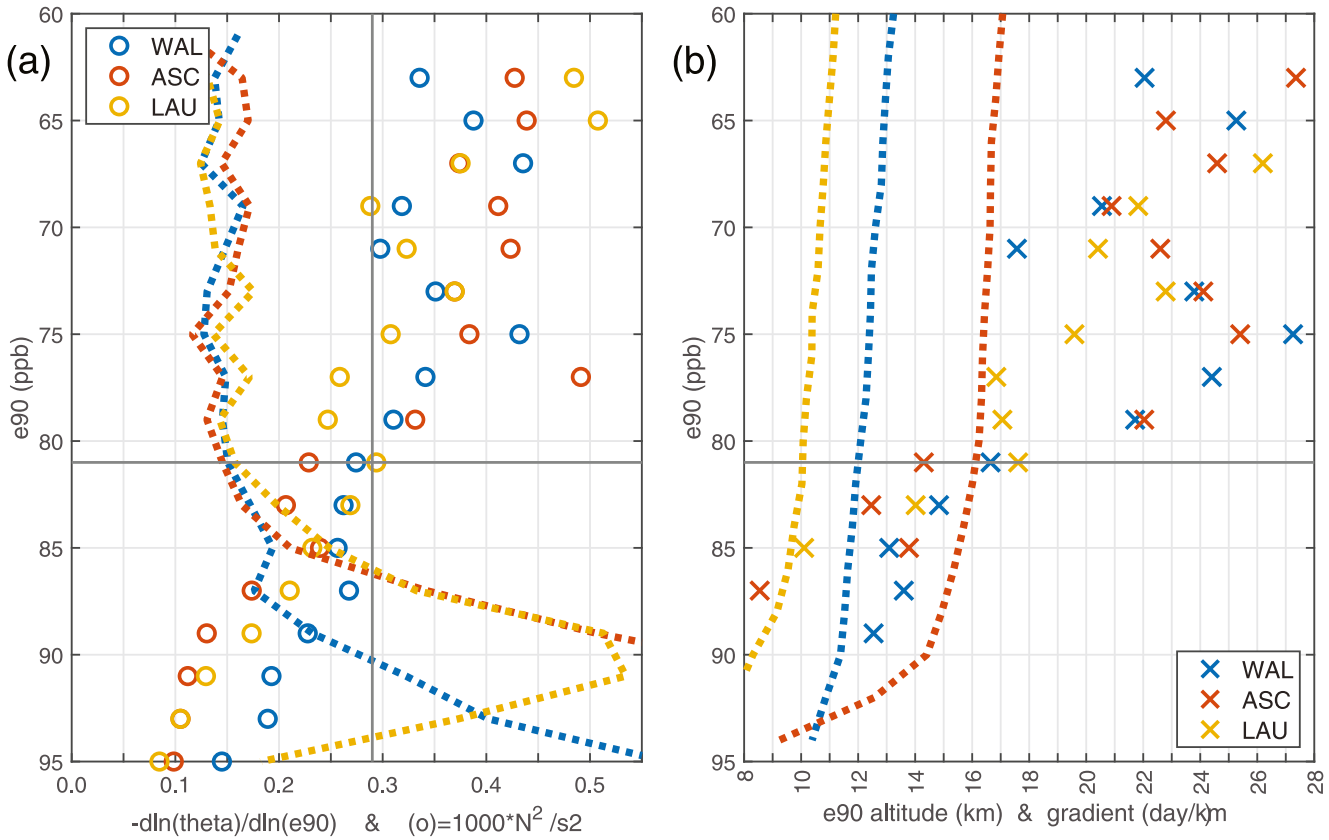

Figure 9. (a) Annual mean of d (ln(θ))/d (ln (e90)) for all 3 sonde sites (colored dashed lines, see legend). Also shown is the Brunt-Väisälä frequency squared (N^2) and scaled by 1000 (colored o's). The vertical gray line marks the stratosphere-troposphere boundary of Tinney et al. (2022) at $N = 0.017\ s^{-1}$. The horizontal gray line marks our best fit of the e90 tropopause threshold at 81 ppb. (b) Annual mean of the altitude (km) of the e90 surface (colored dashed lines) plus the vertical gradient of the e90 age-of-air in days per km defined as $G = 90$ days(ln (e90))/d(z).

methods must be used. The relatively coarse vertical resolution of both models and satellite observations results in a high bias for tpp O_3 that can be estimated from sonde data. Trying to determine tpp O_3 near the jets and over the poles using a gradient method for the tropopause fails (Figure 8) because the transport barriers are horizontal. Fortunately, considering these difficulties, we find that the UCI CTM tpp O_3 values (at e90 = 81ppb) are reasonably close to the sonde data at our three sites (Figure 7).

With this hybrid approach—e90 calibration and sonde verification—to evaluating our models, there are additional areas to explore. The tropopause transition layers seen in observations can be modeled and diagnosed with e90 in a statistical sampling, even with relatively coarse (~0.5 km) vertical model layers. The direct link of e90 or other age-of-air tracer provides a very interesting model diagnostic for O_3 in the lower stratosphere as it quantifies the net rate of change from chemistry and transport. Given the relationship between θ and e90, it is possible that model metrics can be developed for the O_3−θ curves based on observations. Our typical model diagnostics of O_3 near the tropopause cannot be based on monthly mean data and need a refresh.

Conflict of Interest

The author declares no conflicts of interest relevant to this study.

Data Availability Statement

The unique data used in this study are archived as netcdf files along with the plotted data in Prather (2025). The publicly available observations are listed below.

Ozonesondes	Downloaded: .dat's 2023-10-03 (WAL and ASC); csv's 2023-01-08 (LAU) https://tropo.gsfc.nasa.gov/shadoz/Wallops.html, see Witte et al., 2019. https://tropo.gsfc.nasa.gov/shadoz/Ascension.html, see Witte et al., 2017, 2018; Thompson et al., 2017. https://woudc.org/data/explore.php, select "OzoneSonde", "New Zealand (NZL)", and Lauder (256)', "Ecc"
OMPS	downloaded: 2023-04-01 via wget, https://data.gesdisc.earthdata.nasa.gov/data/SNPP_OMPS_Level2/… accepted profiles: "O3Convergence" <10; "O3Status" = 2:7; "O3Quality" = 0; "QMV" = 0; "ASI_PMCFlag" = 0. Processing: "TropopauseAltitude" reported using GMAO FP-IT T profiles
ACE-FTS	downloaded: 2023-12-17 from doi:10.20383/101.0291 https://databace.scisat.ca/level2/ processing: grid is 1 km altitude; drop all $O_3 < 0$ and below 6 km ($k = 1{:}6$). For tropopause, find lowest troposphere level (LR > 2 K/km) with 2 stratosphere levels (LR ≤ 2) above it

Acknowledgments
This research at UCI was supported by National Aeronautics and Space Administration Grant 80NSSC21K1454 and National Science Foundation grant AGS-2135749.

References

Bernath, P., Steffen, J., Crouse, J., & Boone, C. (2020). Atmospheric chemistry experiment SciSat level 2 processed data, v4.0. *Federated Research Data Repository*. https://doi.org/10.20383/101.0291

Bernath, P. F. (2023). The atmospheric chemistry experiment (ACE): Aerosol and gas analysis from orbit. *Trends in Analytical Chemistry*, *166*, 117207. https://doi.org/10.1016/j.trac.2023.117207

Bethan, S., Vaughan, G., & Reid, S. J. (1996). A comparison of ozone and thermal tropopause heights and the impact of tropopause definition on quantifying the ozone content of the troposphere. *Quarterly Journal of the Royal Meteorological Society*, *122*(532), 929–944. https://doi.org/10.1002/qj.49712253207

Birner, T. (2006). Fine-scale structure of the extratropical tropopause region. *Journal of Geophysical Research*, *111*(D4), D04104. https://doi.org/10.1029/2005JD006301

Hegglin, M. I., Boone, C. D., Manney, G. L., & Walker, K. A. (2009). A global view of the extratropical tropopause transition layer from Atmospheric Chemistry Experiment Fourier Transform Spectrometer O_3, H_2O, and CO. *Journal of Geophysical Research*, *114*(D7), D00B11. https://doi.org/10.1029/2008JD009984

Hoffmann, L., & Spang, R. (2022). An assessment of tropopause characteristics of the ERA5 and ERA-Interim meteorological reanalyses. *Atmospheric Chemistry and Physics*, *22*(6), 4019–4046. https://doi.org/10.5194/acp-22-4019-2022

Homeyer, C. R., Bowman, K. P., & Pan, L. L. (2010). Extratropical tropopause transition layer characteristics from high-resolution sounding data. *Journal of Geophysical Research*, *115*(D13), D13108. https://doi.org/10.1029/2009JD013664

Kramarova, N. A., Xu, P., Mok, J., Bhartia, P. K., Jaross, G., Moy, L., et al. (2024). Decade-long ozone profile record from Suomi NPP OMPS Limb Profiler:Assessment of version 2.6 data. *Earth and Space Science*, *11*(9), e2024EA003707. https://doi.org/10.1029/2024EA003707

Lauritzen, P. H., Ullrich, P. A., Jablonowski, C., Bosler, P. A., Calhoun, D., Conley, A. J., et al. (2014). A standard test case suite for two-dimensional linear transport on the sphere: Results from a collection of state-of-the-art schemes. *Geoscientific Model Development*, *7*(1), 105–145. https://doi.org/10.5194/gmd-7-105-2014

Logan, J. A. (1999). An analysis of ozonesonde data for the troposphere: Recommendations for testing 3-D models, and development of a gridded climatology for tropospheric ozone. *Journal of Geophysical Research*, *104*(D13), 16115–16149. https://doi.org/10.1029/1998JD100096

McPeters, R. D., Labow, G. J., & Logan, J. A. (2007). Ozone climatological profiles for satellite retrieval algorithms. *Journal of Geophysical Research*, *112*(D5), D05308. https://doi.org/10.1029/2005JD006823

Pan, L. L., Paulik, L. C., Honomichl, S. B., Munchak, L. A., Bian, J., Selkirk, H. B., & Vömel, H. (2014). Identification of the tropical tropopause transition layer using the ozone-water vapor relationship. *Journal of Geophysical Research: Atmospheres*, *119*(6), 3586–3599. https://doi.org/10.1002/2013JD020558

Prather, M. J. (1986). Numerical advection by conservation of second-order moments. *Journal of Geophysical Research*, *91*(D6), 6671–6681. https://doi.org/10.1029/JD091iD06p06671

Prather, M. J. (2025). Atmospheric data used for calibrating the tropopause in global chemistry-climate or chemistry-transport models [Dataset]. *Dryad*. https://doi.org/10.5061/dryad.pk0p2ngzt

Prather, M. J., & Zhu, X. (2024). Lifetimes and timescales of tropospheric ozone. *Elementa: Science of the Anthropocene*, *12*(1), 00112. https://doi.org/10.1525/elementa.2023.00112

Prather, M. J., Zhu, X., Tang, Q., Hsu, J., & Neu, J. L. (2011). An atmospheric chemist in search of the tropopause. *Journal of Geophysical Research*, *116*(D4), D04306. https://doi.org/10.1029/2010JD014939

Reichler, T., Dameris, M., & Sausen, R. (2003). Determining the tropopause height from gridded data. *Geophysical Research Letters*, *30*(20), 2042. https://doi.org/10.1029/2003GL018240

Santer, B. D., Wehner, M. F., Wigley, T. M. L., Sausen, R., Meehl, G. A., Taylor, K. E., et al. (2003). Contributions of anthropogenic and natural forcing to recent tropopause height changes. *Science*, *301*(5632), 479–483. https://doi.org/10.1126/science.1084123

Schoeberl, M. R. (2004). Extratropical stratosphere-troposphere mass exchange. *Journal of Geophysical Research*, *109*(D13). https://doi.org/10.1029/2004jd004525

Seidel, D. J., & Randel, W. J. (2006). Variability and trends in the global tropopause estimated from radiosonde data. *Journal of Geophysical Research*, *111*(D21), D21101. https://doi.org/10.1029/2006JD007363

Siler, N., & Durran, D. (2015). Assessing the impact of the tropopause on mountain waves and orographic precipitation using linear theory and numerical simulations. *Journal of the Atmospheric Sciences*, *72*(2), 803–820. https://doi.org/10.1175/JAS-D-14-0200.1

Skeie, R. B., Myhre, G., Hodnebrog, Ø., Cameron-Smith, P. J., Deushi, M., Hegglin, M. I., et al. (2020). Historical total ozone radiative forcing derived from CMIP6 simulations. *Npj Climate and Atmospheric Science*, *3*(1), 32. https://doi.org/10.1038/s41612-020-00131-0

Tegtmeier, S., Anstey, J., Davis, S., Dragani, R., Harada, Y., Ivanciu, I., et al. (2020). Temperature and tropopause characteristics from reanalyses data in the tropical tropopause layer. *Atmospheric Chemistry and Physics*, *20*(2), 753–770. https://doi.org/10.5194/acp-20-753-2020

Thompson, A. M., Stauffer, R. M., Wargan, K., Witte, J. C., Kollonige, D. E., & Ziemke, J. R. (2021). Regional and seasonal trends in tropical ozone from SHADOZ profiles: Reference for models and satellite products. *Journal of Geophysical Research: Atmospheres*, *126*(22), e2021JD034691. https://doi.org/10.1029/2021JD034691

Thompson, A. M., Witte, J. C., Sterling, C., Jordan, A., Johnson, B. J., Oltmans, S. J., et al. (2017). First reprocessing of southern hemisphere additional ozonesondes (SHADOZ) ozone profiles (1998-2016): 2. Comparisons with satellites and ground-based instruments. *Journal of Geophysical Research: Atmospheres*, *122*(23), 13000–13025. https://doi.org/10.1002/2017JD027406

Tinney, E. M., Homeyer, C. R., Elizalde, L., Hurst, D. F., Thompson, A. M., Stauffer, R. M., et al. (2022). A modern approach to a stability-based definition of the tropopause. *Monthly Weather Review*, *150*(12), 3151–3174. https://doi.org/10.1175/MWR-D-22-0174.1

Turhal, K., Plöger, F., Clemens, J., Birner, T., Weyland, F., Konopka, P., & Hoor, P. (2024). Variability and trends in the potential vorticity (PV)-gradient dynamical tropopause. *Atmospheric Chemistry and Physics*, *24*(23), 13653–13679. https://doi.org/10.5194/acp-24-13653-2024

Wirth, V. (2000). Thermal versus dynamical tropopause in upper tropospheric balanced flow anomalies. *Quarterly Journal of the Royal Meteorological Society*, *126*(562), 299–317. https://doi.org/10.1002/qj.49712656215

Witte, J. C., Thompson, A. M., Schmidlin, F. J., Northam, E. T., Wolff, K. R., & Brothers, G. B. (2019). The NASA Wallops Flight Facility digital ozonesonde record: Reprocessing, uncertainties, and dual launches. *Journal of Geophysical Research: Atmospheres*, *124*(6), 3565–3582. https://doi.org/10.1029/2018JD030098

Witte, J. C., Thompson, A. M., Smit, H. G. J., Fujiwara, M., Posny, F., Coetzee, G. J. R., et al. (2017). First reprocessing of southern hemisphere ADditional OZonesondes (SHADOZ) profile records (1998-2015): 1. Methodology and evaluation. *Journal of Geophysical Research: Atmospheres*, *122*(12), 6611–6636. https://doi.org/10.1002/2016JD026403

Witte, J. C., Thompson, A. M., Smit, H. G. J., Vömel, H., Posny, F., & Stübi, R. (2018). First reprocessing of southern hemisphere ADditional OZonesondes profile records: 3. Uncertainty in ozone profile and total column. *Journal of Geophysical Research: Atmospheres*, *123*(6), 3243–3268. https://doi.org/10.1002/2017JD027791

Worthington, R. M. (1998). Tropopausal turbulence caused by the breaking of mountain waves. *Journal of Atmospheric and Solar-Terrestrial Physics*, *60*(16), 1543–1547. https://doi.org/10.1016/S1364-6826(98)00105-9

Xian, T., & Homeyer, C. R. (2019). Global tropopause altitudes in radiosondes and reanalyses. *Atmospheric Chemistry and Physics*, *19*(8), 5661–5678. https://doi.org/10.5194/acp-19-5661-2019

Supporting Information for 2025AV001651

Calibrating the tropospheric air and ozone mass

Michael J. Prather, Earth System Science Department, UC Irvine

Search for a universal PV or O_3 tropopause indicator

From **Figures 5a** and **6abc**, we show that an e90 threshold of about 80±5 ppb follows the lapse rate tropopause across the three latitudes of our sonde sites and throughout the year. Thus, the e90 tracer provides a tpp threshold that is mostly independent of season and latitude (i.e., universal). Can we identify a similar universal tpp threshold for other non-synthetic (measurable) atmospheric tracers? Defining the tropopause boundary using dynamical and chemical tracers such as potential vorticity (PV, K m^2 kg^{-1} s^{-1}) and O_3 (ppb) is well established (e.g., Bethan et al., 1996; Wirth, 2000; Birner, 2006, Tinney et al., 2022). On synoptic scales, PV & O_3 follow the tropopause and its deformation because both are conserved quantities; but on monthly time scales, diabatic heating for PV and chemistry for O_3 make these less than reliable tracers.

As demonstrated from ozonesondes (Logan, 1999), the tropopause O_3 abundance varies with latitude and season because on time scales of a month photochemistry alters the abundance, particularly the tropospheric lower boundary condition. Some models adopt a fixed tropopause O_3 threshold (e.g., 150 ppb) for which they can diagnose the O_3 chemical budgets and stratosphere-troposphere exchange (STE) fluxes. Unfortunately, we know this threshold changes with latitude and season and the diagnosed fluxes depend sensitively (~20%) on the choice of threshold (Hsu and Prather, 2014, Table 1). Another problem with an O_3 tropopause is that model chemistries can vary, and the tropopause boundary should not change with the choice of O_3 chemical model, particularly if we are using it for chemistry diagnostics.

For PV, we have a similar problem: it is not conserved under diabetic heating, and diabatic heating like chemistry varies from model to model. Schoeberl (2004) notes that "unfortunately, no standard definition for the extratropical PV tropopause exists" and his incredible analysis of the tropopause height difference between the WMO lapse rate tropopause (LRT) and PV thresholds (his Figure 2) shows that 2 to 4 PVU (1 PVU = 10^{-6} K m^2 kg^{-1} s^{-1}) roughly aligns with the WMO tropopause.

We were wary of calculating the 3D spatial gradients needed for PV from our model's relatively coarse grid (100 km horizontal by 1 km vertical at tropopause); but with the new meteorological fields (IFS Cycle 43r3), the forecast PV for each cell in the forecast fields was saved. Using a three-year simulation with snapshots at the first of each month, we did a parallel investigation of PV and O_3 about the tropopause, searching for spatially and seasonally consistent e90-PV and e90-O_3 relationships that would demonstrate some universal calibration of the tropopause.

Figures S1 and **S2** show scatter plots of PV vs. e90 and O_3 vs. e90, respectively, for six latitude bins from a single snapshot (00H 1 Jan 2005). The latitude bands avoid the core tropics and the poles: 61±5ºS, 41±5ºS, 21±5ºS, 21±5ºN, 41±5ºN, and 61±5ºN. For each figure, we calculate the statistical range, 16^{th}-50^{th}-84^{th} percentile, of PV and O_3 for four e90 bins 60±5 ppb, 70±5 ppb, 80±5 ppb, and 90±5 ppb. For the most part, the median PV matches the mean and the 68% confidence range is two standard deviations (not shown), implying that we have a near-normal distribution. Thus, the PV vs e90 relationship is very well defined. The O_3 vs e90 relationship is similar with the exception of the e90 = 90±5 ppb bin in the subtropics, where e90

reaches far into the troposphere. The percentiles for the e90 bins, including the intermediate altitude belts not shown in Figures S1 & S2 (51±5ºS, 31±5ºS, 31±5ºN, 51±5ºN), are summarized in **Figure S3** for both PV and O_3. For PV, each e90 bin (except for 90 ppb) shows a V-shaped pattern with a minimum in the subtropics. The consistency here shows that PV and e90 are both marking the age-transport barriers in the lower stratosphere but there is a large symmetric latitudinal shift in the PV threshold. This result is similar to Schoeberl's findings. For O_3, there is a near-linear increase of 50 ppb from 61ºS to 61ºN, again indicating that O_3 is a reasonable measure of the tropopause, but one that slides with latitude. One must expect that the O3-e90 calibration will change with other model's chemistry.

The seasonal (monthly) evolution of the PV-e90 and O_3-e90 calibrations for the same latitude belts as in Figures S1-S2 are shown in **Figures S4 & S5**. The seasonal variation in median PV often exceeds 2 PVU and changes with latitude. For O_3, the NH median O3 peaks in spring (61ºN) to late summer (21ºN) driven by tropospheric pollution, with an annual amplitude of more than 100 ppb. In the SH at 41ºS the seasonal cycle is damped as expected from the Lauder sondes, but at 61ºS and 21ºS the cycles are large and of opposite phase.

Unfortunately, we end up reaffirming the need for a complex, latitude-by-month calibration of the tropopause values for PV and O_3. An age tracer like SF6 with a very long lifetime and slowly increasing emissions might provide a clear, measurable tracer definition of the tropopause, but with NH sources, one would have to treat the hemispheric tpp definition differently. A possible problem with a long-lived tracer is that we are trying to measure temporal gradients of order 20 days to resolve 1 km in tpp height (see vertical mixing rates in **Figure 9**). Thus, the precision of in situ sampling would need to be of order 5% of the annual growth rate.

References

Bethan, S., G. Vaughan, and S. J. Reid (1996), A comparison of ozone and thermal tropopause heights and the impact of tropopause definition on quantifying the ozone content of the troposphere, Q. J. R. Meteorol. Soc., 122, 929– 944. doi: 10.1002/qj.49712253207.

Birner, T. (2006), Fine-scale structure of the extratropical tropopause region, J. Geophys. Res., 111, D04104, doi:10.1029/2005JD006301.

Hsu, Juno and M.J. Prather (2014), Is the vertical residual velocity a good proxy for stratosphere-troposphere exchange of ozone? Geophys. Res. Lett., 41, doi:10.1029/2014GL061994.

Logan, J. A. (1999), An analysis of ozonesonde data for the troposphere: Recommendations for testing 3-D models, and development of a gridded climatology for tropospheric ozone, J. Geophys. Res., 104, 16,115 – 16,149, doi: 10.1029/1998JD100096.

Schoeberl, M. R. (2004), Extratropical stratosphere-troposphere mass exchange. Journal of Geophysical Research, 109(D13). doi: 10.1029/2004jd004525.

Tinney, E.M., Homeyer, C.R., Elizalde, L., Hurst, D.F., Thompson, A.M., Stauffer, R.M., Vömel, H., and Selkirk, H.B. (2022), A Modern Approach to a Stability-Based Definition of the Tropopause, Mon.Wea.Rev., 3151–3174, DOI: 10.1175/MWR-D-22-0174.1

Wirth, V. (2000), Thermal versus dynamical tropopause in upper tropospheric balanced flow anomalies, Q. J. R. Meteorol. Soc., 126, 299– 317. doi: 10.1002/qj.49712656215

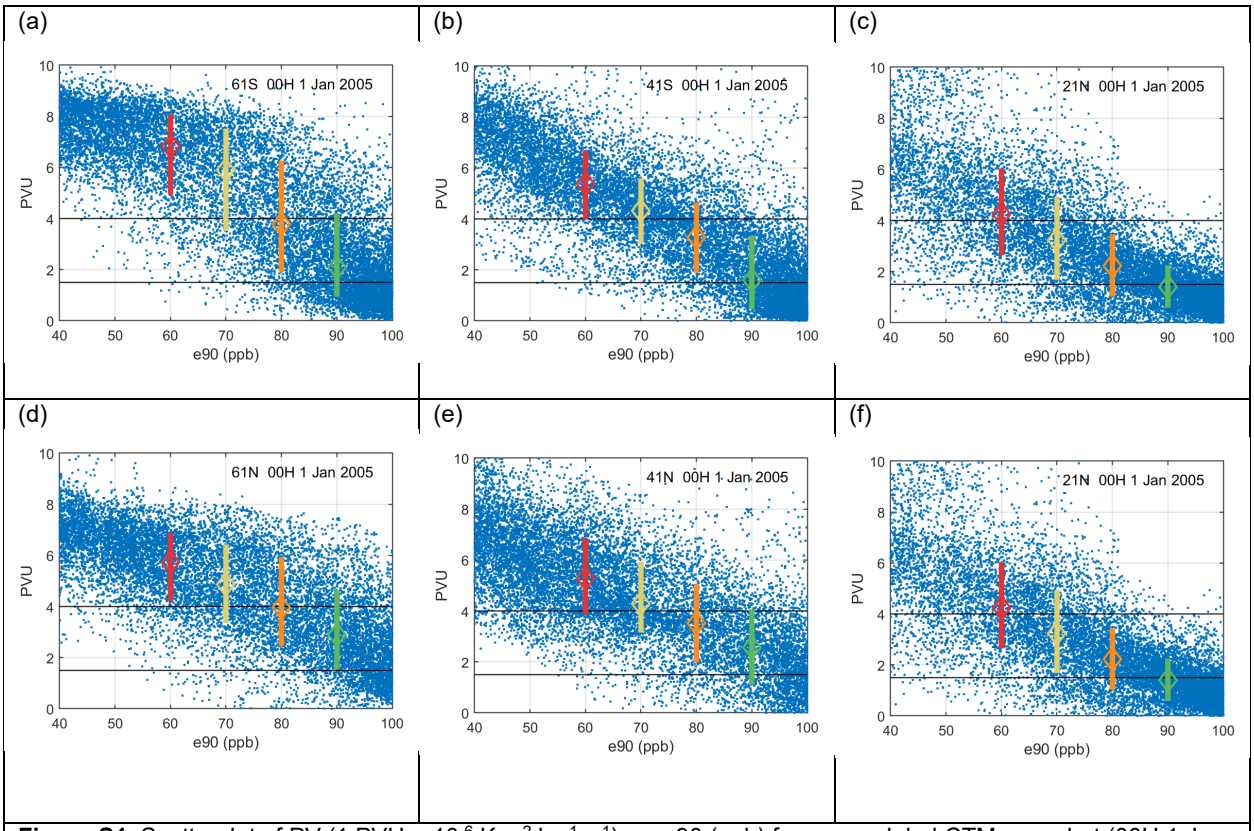

Figure S1. Scatterplot of PV (1 PVU = 10^{-6} K m^2 kg^{-1} s^{-1}) vs. e90 (ppb) from one global CTM snapshot (00H 1 Jan 2005). Typical range of tpp PV (1.5 and 4.0 PVU) shown with horizontal black lines. Each plot samples all cells within ±5º of latitude about the given values: **(a)** 61ºS, **(b)** 41ºS, **(c)** 21ºS, **(d)** 21ºN, **(e)** 41ºN, and **(f)** 61ºN. The colored diamonds and vertical bars show the percentiles (16%, 50%, 84%) of the PVU values sampled for e90 = 60±5 (red), 70±5 (gold), 80±5 (orange), and 90±5 (green) ppb. e90 = 80 ppb (orange) is closest to the tropopause.

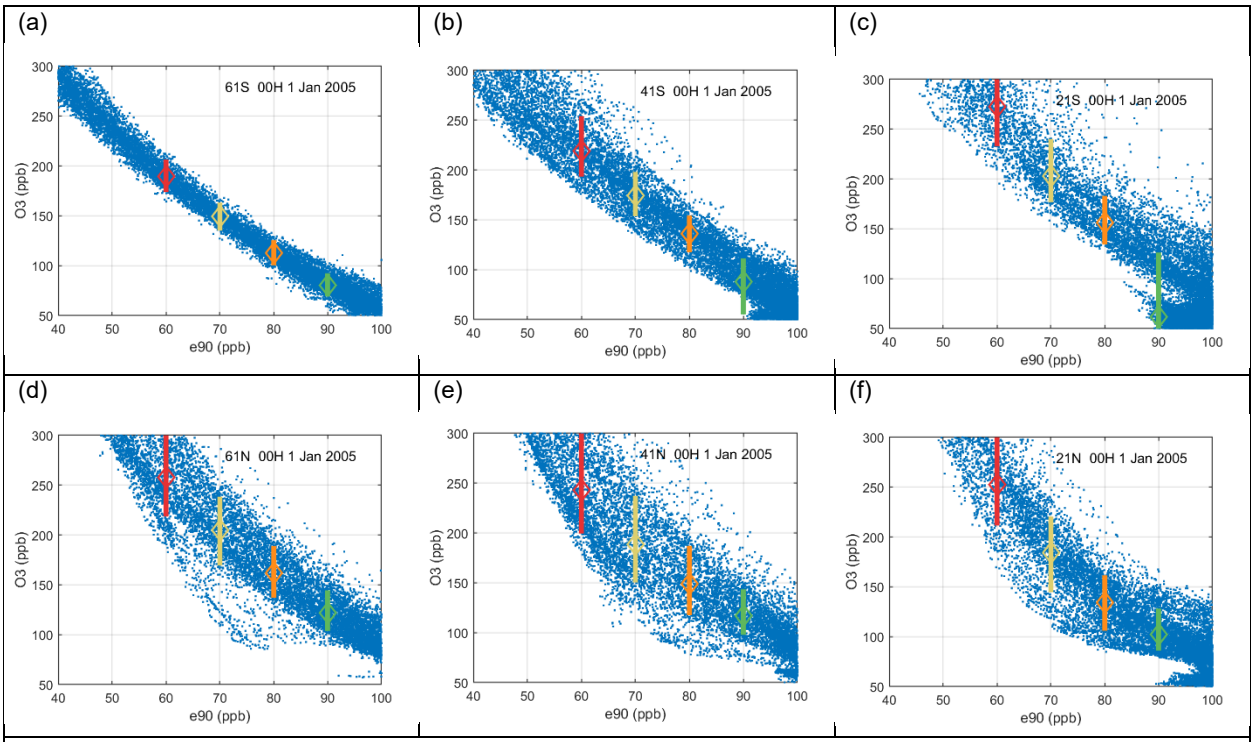

Figure S2. Scatterplot of O_3 (ppb) vs. e90 (ppb) from one global CTM snapshot (00H 1 Jan 2005). See Figure S1.

95

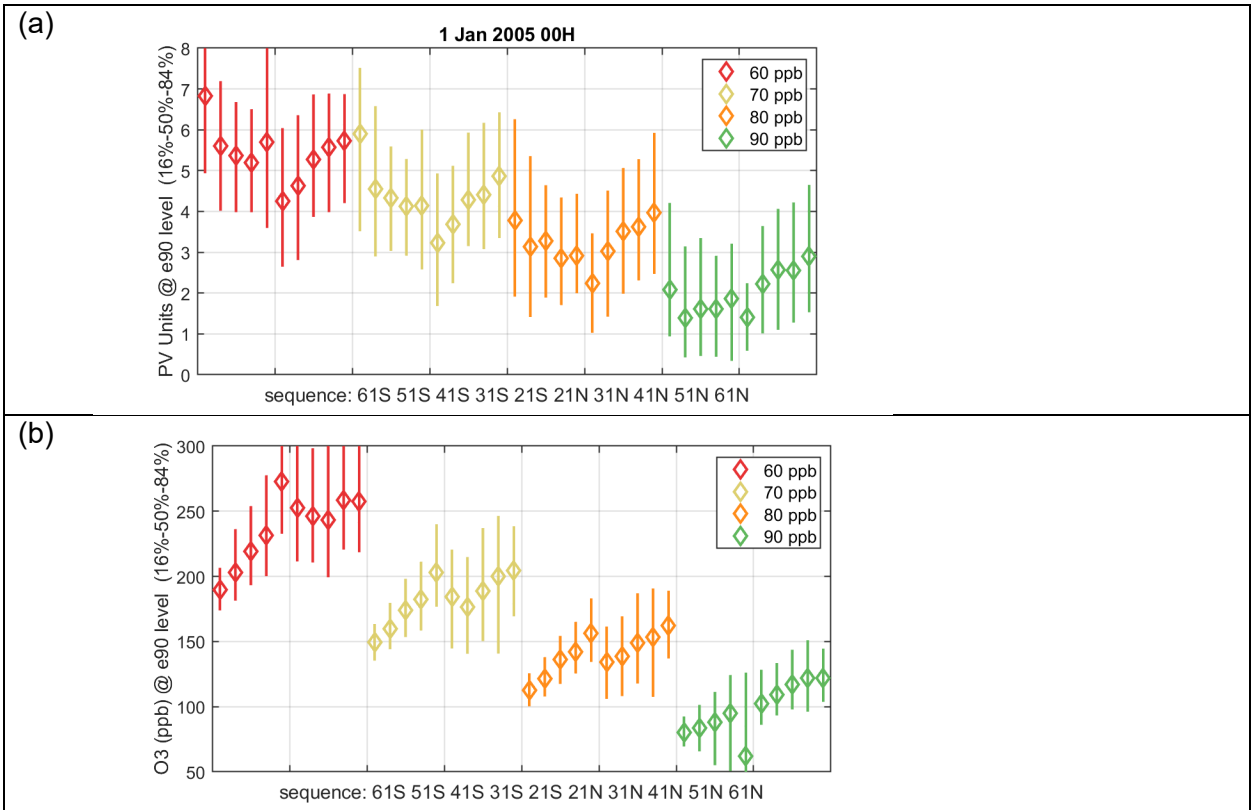

Figure S3. The percentile statistics (16%, 50%, 84%) for **(a)** PV in PVU and **(b)** O_3 in ppb taken from the graphs in Figures S1 & S2, based on a single CTM snapshot at 00H 1 Jan 2005. Each color squence represents a different e90 value from 60 ppb (red, above the tropopause) to 90 ppb (green, often below the tropopause), see legend. The sequence of ten bars for each color is for the ±5º latitude bins about 61ºS, 51ºS, 41ºS, 31ºS, 21ºS, 21ºN, 31ºN, 41ºS, 51ºN, and 61ºN, from left to right.

96
97

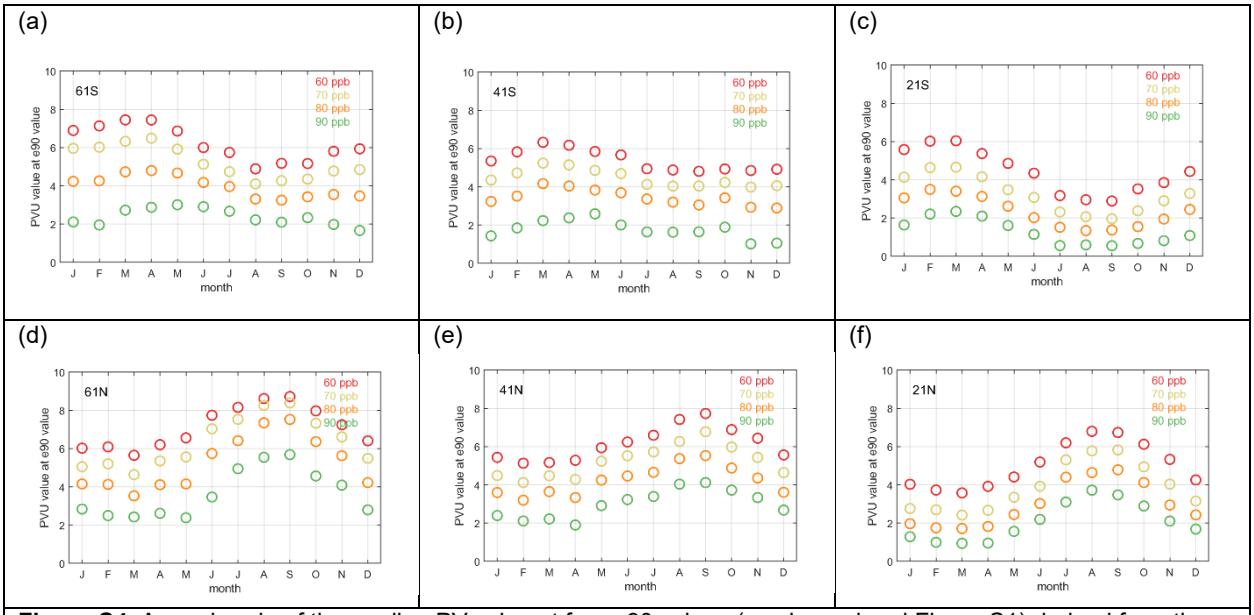

Figure S4. Annual cycle of the median PV value at four e90 values (see legend and Figure S1) derived from the average of three years of monthly snapshots (00H for the 1st of the month from 1 Feb 2002 to 1 Jan 2005). Based on work here, the e90 = 80 ppb (green in this graph) is closest to the tropopause.

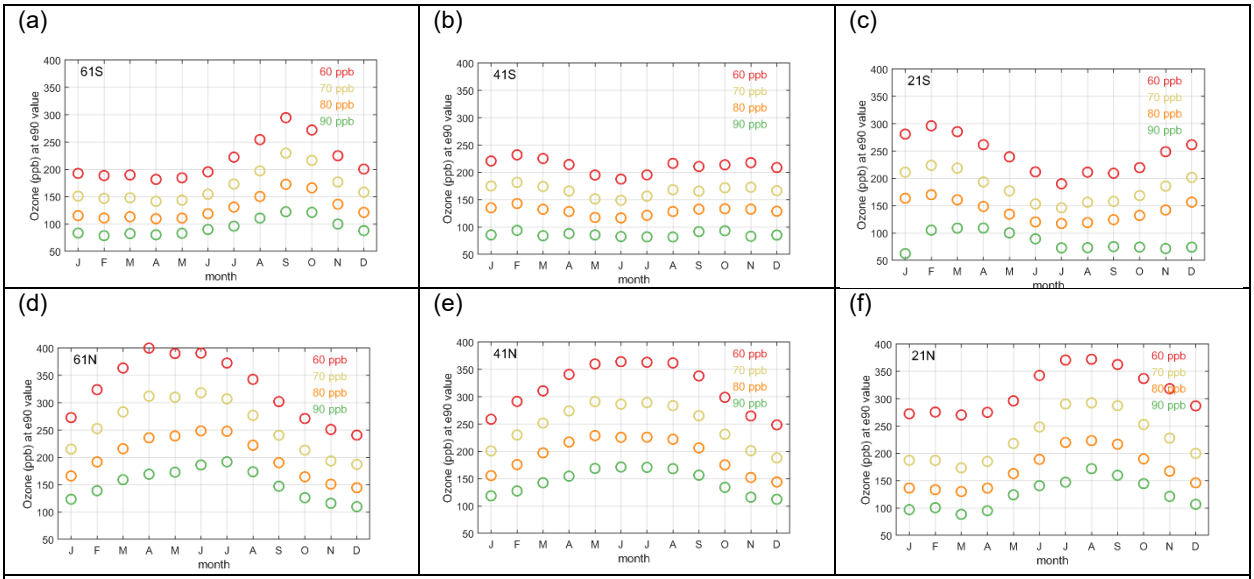

Figure S5. Annual cycle of the median O_3 value at four e90 values (see legend and Figure S1) derived from the average of three years of monthly snapshots (00H for the 1st of the month from 1 Feb 2002 to 1 Jan 2005). See Figure S4.