



1 **An improved Trajectory-mapped Ozonesonde dataset for the Stratosphere and Troposphere**
2 **(TOST): update, validation and applications**

3

4 Zhou Zang¹, Jane Liu¹, David Tarasick², Omid Moeini², Jianchun Bian³, Jinqiang Zhang³, Anne
5 M. Thompson^{4,5}, Roeland Van Malderen⁶, Herman G.J. Smit⁷, Ryan M. Stauffer⁴, Bryan J.
6 Johnson⁸ and Debra E. Kollonige^{4,9}

7 ¹Department of Geography and Planning, University of Toronto, Toronto, Canada

8 ²Environment and Climate Change Canada, Toronto, Canada

9 ³Key Laboratory of Middle Atmosphere and Global Environment Observation, Institute of
10 Atmospheric Physics, Chinese Academy of Sciences, Beijing, China

11 ⁴Atmospheric Chemistry and Dynamics Laboratory, NASA Goddard Space Flight Center,
12 Greenbelt, Maryland, USA

13 ⁵University of Maryland Baltimore County, Baltimore, MD, USA

14 ⁶Royal Meteorological Institute of Belgium, Brussels, Belgium

15 ⁷Institute for Energy and Climate Research: Troposphere (IEK-8), Research Centre Juelich (FZJ),
16 Juelich, Germany.

17 ⁸NOAA/ESRL Global Monitoring Division, Boulder, Colorado, USA

18 ⁹Science Systems and Applications, Inc., Lanham, MD, USA

19

20 *Correspondence: Jane Liu (janejj.liu@utoronto.ca)*

21

22 **Abstract**

23 A global-scale horizontally- and vertically-resolved ozone climatology can provide a detailed
24 assessment of ozone variability. Here, the Trajectory-mapped Ozonesonde dataset for the
25 Stratosphere and Troposphere (TOST) ozone climatology is improved and updated to the recent
26 decade (1970s-2010s) on a grid of $5^\circ \times 5^\circ \times 1$ km (latitude, longitude, and altitude) from the surface
27 to 26 km altitude, with the most recent ozonesonde data re-evaluated following the ASOPOS-2
28 guidelines (GAW Report No. 268, 2021). Comparison between independent ozonesonde and
29 trajectory-derived ozone shows good agreement in each decade, altitude, and station, with relative



30 differences (RD) of 2-4% in the troposphere and 0.5% in the stratosphere. Comparisons of TOST
31 with aircraft and two satellite datasets, the Satellite Aerosol and Gas Experiment (SAGE) and the
32 Microwave Limb Sounder (MLS), show comparable overall agreement. The updated TOST
33 outperforms the previous version with higher data coverage in all latitude bands and altitudes and
34 14-17% lower RD compared to independent ozonesondes, employing twice as many ozonesonde
35 profiles and an updated trajectory simulation model. Higher uncertainties in TOST are where data
36 are sparse, i.e., over the southern high latitudes and the tropics, and before the 1980s, and where
37 variability is high, i.e., at the surface and upper troposphere and lower stratosphere (UTLS).
38 Caution should therefore be taken when using TOST in these spaces and times. TOST captures
39 global ozone distributions and temporal variations, showing an overall insignificant change of
40 stratospheric ozone after 1998. TOST offers users a long record, global coverage, and high vertical
41 resolution.

42

43 **1. Introduction**

44 The global ozone distribution and its long-term changes at different altitudes, longitudes, and
45 latitudes are critical to understanding global ozone variability and its interactions with climate
46 change. While the ozone trends themselves can indicate the impact of changes in climatic
47 dynamics (Hassler et al., 2008), or chemistry, including the effect of the Montreal Protocol
48 (Steinbrecht et al., 2017), long-term horizontally- and vertically-resolved ozone are needed for
49 prescribing, evaluating and refining ozone simulations in climate models (Hassler et al., 2018),
50 and to quantify changes in radiative forcing and projecting reliable future climate scenarios
51 (Nowack et al., 2015).

52 Balloon-borne ozonesondes are the principal source of trend-quality long-term records of
53 ozone profiles below ~18 km (Tarasick et al., 2021). However, the horizontal and temporal



coverages of ozonesondes are limited by the sparse distribution of the stations (less than 100 worldwide) and their low launch frequency (1-3 times/week) (Liu et al., 2013a). The In-Service Aircraft for a Global Observing System (IAGOS) program has measured ozone profiles worldwide since 1994 via the instruments onboard a number of commercial aircraft, with high sampling frequency at some airports (Thouret et al., 1998). However, sampling is unevenly distributed both spatially and temporally because the flights are constrained by commercial airlines' operation schedules. Satellite observations have the advantage of providing large-scale 3-dimensional ozone data with consistent quality. However, satellite data are provided for the stratosphere only or for troposphere with limited vertical resolution (6–10 km) (Worden et al., 2007; Liu et al., 2010; Tarasick et al., 2019b) caused by uncertainties for satellites to retrieve tropospheric ozone through the large stratospheric ozone burden (Bhartia, 2002). A number of studies have developed long-term (since the 1980s) ozone climatologies by combining ozone data from ozonesondes and/or multiple satellite instruments (McPeters et al., 2007; MCPeters and Labow, 2012; Hassler et al., 2018; Bodeker et al., 2021; Bognar et al., 2022), but these are generally zonally-averaged. Chemistry–climate models are also used to develop 3-dimensional ozone data fields, especially for long-term, global-scale simulations (Eyring et al., 2010; Chen et al., 2018); these models present our best understanding of processes controlling ozone variations but still suffer from large uncertainties regarding the inventories, parameterizations, radiation transport schemes, and simulation of the atmospheric circulations and systems (Young et al., 2018; Wild et al., 2020; Griffiths et al., 2021; Zeng et al., 2022).

Liu et al. (2013a, b) constructed a long-term 3-dimensional global-scale ozone dataset using a trajectory-mapping method, extending sparse ozonesonde measurements and filling gaps in the spatial domain by backward and forward trajectory simulations. The trajectory-mapping method assumes the ozone mixing ratio in the same air parcel along each trajectory path is constant for



several days, which is reasonable given that the lifetime of ozone in most of the troposphere and stratosphere ranges from weeks to months (Jacob, 1999). The result is a global dataset that is independent of satellite measurements and photochemical modeling processes. The trajectory-mapping method can be characterized as a meteorologically-guided interpolation method, which necessarily carries more information than conventional statistical interpolation methods (Stohl et al., 2001). In addition, the trajectory-derived ozone data cover higher latitudes (to 90°N and 90°S) and a longer time period (since the 1960s) (Liu et al., 2013b). The Trajectory-mapped Ozone-sonde dataset for the Stratosphere and Troposphere (TOST) Version 1 is available from 1965-2012 at the World Ozone and UV Data Centre (WOUDC, <https://woudc.org/archive/products/ozone/vertical-ozone-profile/ozonesonde/1.0/tost/>, last access: Jan 29, 2024), and has been successfully applied in model evaluation (Skeie et al., 2020; Badia et al., 2021), ozone and climate trend studies (Polvani et al., 2017; Gaudel et al., 2018; Gulev et al., 2021), as a background ozone climatology (Xu et al., 2018; Moeini et al., 2020, and for tropospheric ozone burden estimation (Griffiths et al., 2021).

There have been several important developments since the publication of the first version of TOST data in 2013 (Liu et al. 2013a, b), which we refer to as TOST Version 1, or TOST-v1. An improved version of TOST, namely TOST-v2, is necessary for the following reasons. Firstly, there are some 50,000 new ozone profiles, many from newly established ozonesonde stations (see Section 2.1). These new ozonesonde data permit updating TOST to 2021, providing 3-dimensional ozone information through the 2010s. Secondly, data from many ozonesonde stations have been updated to higher-quality versions. An important source of uncertainty in TOST-v1 is possible biases in station records due to instrument changes and/or changes in operating procedures. Homogenized time series are now available from the Harmonization and Evaluation of Ground Based Instruments for Free Tropospheric Ozone Measurements (HEGIFTOM) project for over 40



102 ozonesonde stations (Table S1). For these records, biases due to instrument changes, sensing
103 solution, and preparation changes have been corrected, to reduce the overall uncertainty from 10-
104 20% to 5-10% (Smit and Thompson, 2021). This effort to improve data quality also uncovered an
105 apparent change of bias at stations flying one type of sonde (Stauffer et al., 2020; 2022); 14 global
106 ozonesonde stations (the bolded stations in Table S1) have shown an apparent drop-off of 2-4 %
107 in stratospheric ozone and total ozone column since circa 2013, due to a possible instrument artifact.
108 This is the subject of ongoing research (e.g. [https://gml.noaa.gov/annualconference/abstracts/78-](https://gml.noaa.gov/annualconference/abstracts/78-230424-A.pdf)
109 [230424-A.pdf](https://gml.noaa.gov/annualconference/abstracts/78-230424-A.pdf), last access: Jan 29, 2024). For these stations, ozone measurements above 40 hPa
110 (~20 km) are not recommended for trend calculations. We need therefore to exclude data above
111 40 hPa for the affected profiles in constructing TOST. Thirdly, the version 4.9 of the Hybrid Single-
112 Particle Lagrangian Integrated Trajectory (HYSPLIT) model (Draxler and Hess, 1998) used for
113 trajectory simulation has been improved and updated to version 5.2. Here, we address the
114 mentioned issues and construct an improved and updated TOST using the most state-of-the-art
115 HYSPLIT and most updated ozonesonde data. While Liu et al. (2013a, b) validated TOST-v1 with
116 ozonesonde data at 20 selected stations, TOST-v2 is validated against the ozonesonde data at all
117 141 stations individually with the trajectory-mapped approach omitting the input from the station
118 being tested. In addition, comparisons are made with the IAGOS measurements in the troposphere,
119 and with two limb-viewing satellite sensors, the Satellite Aerosol and Gas Experiment (SAGE)
120 and the Microwave Limb Sounder (MLS), in the stratosphere. This more comprehensive validation
121 and associated uncertainty analysis demonstrates the improved quality of TOST-v2, and also
122 provides some caveats for users of TOST.

123 In the following, Section 2 explains the data sources and the improved trajectory-mapping
124 methodology. Section 3 presents independent validations, comparisons with satellite data, and
125 improvements compared to TOST-v1, as well as the uncertainties in TOST-v2. Based on TOST-



v2, we characterize global ozone variations in the troposphere and stratosphere, and show stagnant stratospheric ozone variation since the late 1990s in Section 4, followed by conclusions in Section 5.

2. Data and Methods

2.1 Ozone sonde data

Ozone sonde data over 1970-2021 at 141 ozone sonde stations worldwide (Figure 1) were downloaded from the World Ozone and Ultraviolet Radiation Data Centre (WOUDC, https://woudc.org/archive/Archive-NewFormat/OzoneSonde_1.0_1/), or where available, homogenized data from Southern Hemisphere Additional Ozone sondes (SHADOZ, <https://doi.org/10.57721/SHADOZ-V06>, last access: Jan 29, 2024) and HEGIFTOM (<https://hegiftom.meteo.be/datasets/ozone sondes>, last access: Jan 29, 2024). The homogenized ozone sonde stations from HEGIFTOM include ozone sonde stations from the SHADOZ network (Thompson et al., 2017; Witte et al., 2017; 2018), the Canadian network (Tarasick et al., 2016), the US network (Sterling et al., 2018), the Network for the Detection of Atmospheric Composition Change (NDACC) and several individual stations (Van Malderen et al., 2016; Witte et al., 2019; Ancellet et al., 2022), with an overall accuracy of 3-5% in both the stratosphere and troposphere. Ozone sonde data from the Beijing Nanjiao Meteorological Observatory (116.47°E, 39.81°N) in Beijing, China, are provided by the Institute of Atmospheric Physics (IAP), Chinese Academy of Sciences. The ozone profiles at Beijing are measured by the Brewer-Mast type GPSO3 ozone sonde and the IAP electrochemical concentration cell (ECC) ozone sonde, which are in fair agreement with commercial ECC ozone sondes (Wang et al., 2003; Xuan et al., 2004; Bian et al., 2007) in both laboratory and field experiments (Zhang et al., 2021; Zeng et al., 2023). In total, data from 43 more stations were used in this version of TOST than in TOST-v1 (Liu et al., 2013b).



150 Figure 1a provides an overview of the distribution of the ozonesonde stations, the number of
151 profiles, and the beginning year for every station. Most of the stations with data before the 1980s
152 are located in North America, Europe, and East Asia. The majority of the stations in the Southern
153 Hemisphere start measurement in the 1990s or later, and so the Southern Hemisphere contains a
154 smaller number of ozone profiles than in the Northern Hemisphere. Figure 1b shows that the total
155 number of ozonesonde profiles per year has almost doubled since the 1990s and reached a
156 maximum in the late 2000s with over 3000 profiles per year. Since then, the available amount of
157 ozonesonde profiles has declined slightly to 2000-3000 profiles per year. The average annual
158 number of profiles per station slightly increased since the 1990s and has stabilized at about 40
159 profiles per year.

160 All the ozonesonde profiles were processed into 1-km vertical resolution by integrating and
161 averaging the ozone volume mixing ratio in 1-km layers from the ground level. The ozonesonde
162 data above 26 km were excluded as the data above this height show large uncertainties at mid- and
163 high-latitudes (Fioletov et al., 2006).

164

165 **2.2 Trajectory simulation**

166 Forward and backward trajectories in four days were calculated every 6 hours using the
167 version 5.2 HYSPLIT model (Stein et al., 2015). HYSPLIT was driven by the reanalysis of hourly
168 meteorological data from the National Centers for Environmental Prediction/National Center for
169 Atmospheric Research (NCEP/NCAR), which has a horizontal resolution of 2.5° by 2.5° in latitude
170 and longitude and 17 vertical levels from the surface to 10 hPa (Kalney et al., 1996). The length
171 of the trajectories influences the spatial coverage and accuracy of the ozone mapping. Generally,
172 uncertainties increase rapidly along the trajectories, with typical errors of about 100–200 km day⁻¹
173 (Stohl, 1998). Trajectories have horizontal uncertainties of 350–400 km after 3 days and 600-1000



174 km after 4 days in the Northern Hemisphere (Engström and Magnusson, 2009). Trajectories show
175 typical vertical deviations of about 200, 800, and 1000 m after 2, 4, and 6 days in the stratosphere,
176 and even greater uncertainties in the troposphere (Stohl and Seibert, 1998). Therefore, to limit
177 trajectory errors, 4-day trajectories were used herein, following previous studies (Tarasick et al.,
178 2010; Liu et al., 2013 a, b).

179

180 **2.3 Three-dimensional ozone mapping based on ozonesonde profiles and trajectories**

181 Ozone mixing ratios from each sounding at the 26 levels were assigned to the corresponding
182 forward and backward trajectory paths. These ozone values at positions every 6 hours along the 4-
183 day backward and forward trajectories (32 positions for each level) were averaged in bins of 5°
184 latitude and 5° longitude, for each 1-km altitude for every month. This bin size corresponds both
185 to the typical uncertainties of 4-day trajectories discussed above, and to the typical ozone
186 correlation length (500-1500 km) in the troposphere and the stratosphere (Liu et al., 2009).
187 Ozonesonde profile data and trajectories in both the troposphere and stratosphere were used to
188 represent the exchanges of ozone between the troposphere and stratosphere in ozone climatology.
189 Based on this mapping, TOST was generated at 26 altitude levels in monthly means for each
190 decade from the 1970s to the 2010s and in annual means for each year from 1970 to 2021.

191 Errors in the mapped data can come from trajectory errors, and from ignoring ozone chemistry
192 (production and loss) along the transport pathway and deposition in the surface layer (Liu et al.,
193 2013a). Differences between the results of backward and forward trajectory mapping can provide
194 a measure of these errors, since in the absence of such errors the results of forward-only and
195 backward-only trajectory mapping should be identical. Therefore, mappings from the forward-
196 only and backward-only trajectories were compared as an initial quality check. Figure S1 shows
197 monthly means (January and July) in 2000 at 3-4 km and 19-20 km, for forward-only and



backward-only mapping. In general, the differences between the two mappings are commonly less than 15% and have no distinct pattern, indicating that trajectory errors, and those from ozone chemistry and deposition, are not systematic. These modest differences between forward-only and backward-only trajectory-mapped ozone fields also validate the reliability of this trajectory-mapping method; both backward and forward trajectories, therefore, were combined in TOST to achieve better averages and higher spatial coverage.

The resulting ozone fields are given in two altitude coordinates (altitude above sea level and altitude above ground level) for users' convenience. In addition, three ozone climatology datasets are generated based on trajectories from ozonesonde observations in both the troposphere and stratosphere, trajectories from observations only in the troposphere (troposphere-only) and trajectories from observations only in the stratosphere (stratosphere-only). Examples presented in this paper all use ozone mapping based on trajectories from observations in both the troposphere and stratosphere, with altitudes above sea level. For this coordinate system, both ozonesonde profiles and mapped data necessarily begin at the altitude of the surface, leaving the levels below as null.

2.4 Validations of TOST

To comprehensively validate TOST, several validations and comparisons were conducted.

2.4.1 Ozonesonde profiles for validation

The first method is to compare the actual ozone profile at each of the ozonesonde stations with the trajectory-derived ozone profile for that station without the input of that station itself. This method is computationally intensive, as the trajectory mapping must be re-calculated (with data for all stations except one), for each ozonesonde station, but it directly tests the reliability of deriving ozone concentrations at a location by integrating the contributions via trajectories from



222 surrounding sites, which is the essential assumption of the trajectory-mapping method. We refer
223 to this set of data that selectively excludes the local data at each station as “Traj-derived”.

224 **2.4.2 Satellite ozone profile data**

225 TOST is further compared with two well-known satellite limb sounder datasets, the Satellite
226 Aerosol and Gas Experiment (SAGE) and the Microwave Limb Sounder (MLS).

227 SAGE II was launched into a 57-degree inclination orbit on board Earth Radiation Budget Satellite
228 (ERBS), and was in operation from 1984–2005. Using the highly accurate solar occultation
229 technique, SAGE can resolve layers in the middle and upper troposphere at 1-km vertical
230 resolution (Kent et al., 1993), with the highest accuracy over the 20–45 altitude range (Cunnold et
231 al., 1996). Here we use the Version 7.0 SAGE II ozone mixing ratio
232 (<https://sage.nasa.gov/missions/about-sage-ii/>, last access: Jan 29, 2024) in the 1980s and 1990s
233 for the comparison.

234 The MLS, onboard the Aura satellite, can measure stratospheric ozone profiles with a vertical
235 resolution of about 3 km. MLS observes microwave radiances that are both emitted and absorbed
236 by the atmosphere. The retrieval is more complex, and uses the optimal estimation approach. Here
237 we use the Version 5.0 MLS ozone mixing ratio
238 (https://disc.gsfc.nasa.gov/datasets/ML2O3_005/summary?keywords=ML2O3_005, last access:
239 Jan 29, 2024) in the 2000s and 2010s for the comparison.

240 **2.4.3 Aircraft ozone profile data**

241 The IAGOS network (<https://www.iagos.org/>, last access: Jan 29, 2024) has been measuring ozone
242 profiles worldwide since 1994 via dual-beam ultraviolet absorption monitors onboard commercial
243 aircraft (Petzold et al., 2015), with an accuracy of about $\pm (2 \text{ nmol mol}^{-1} + 2\%)$ (Nédélec et al.,
244 2016). Ozone monitors are calibrated annually to a reference analyser at the Bureau Internationale
245 des Poids et Mesures (BIPM), and also compared every 2 hours to an in-flight ozone calibration



246 source. Generally good agreement is found between IAGOS profiles and ozonesondes, with
247 positive biases for the sondes of 5-10% (Tilmes et al., 2012; Zbinden et al., 2013; Staufer et al.,
248 2013, 2014; Tanimoto et al., 2015; Tarasick et al., 2019b), making IAGOS ozone suitable for the
249 validation of TOST. Here, the IAGOS ozone profiles were processed into 1 km layers from sea
250 level and matched with the TOST ozone for each level to examine the performance of TOST in
251 the troposphere.

252

253 **3. Validations and comparisons of TOST**

254 **3.1 Validations with ozonesonde observations**

255 First, we show the overall comparison in monthly mean ozone profile between ozonesonde and
256 trajectory-derived values without the inputs of the stations being tested (Traj-Derived), from all
257 the existing stations at selected altitude levels. Note that the full TOST dataset would be better
258 than “Traj-Derived ozone”, especially at the sampling locations because the input of the local
259 station is included in the full TOST data. The three altitude levels are selected to present the overall
260 accuracy of TOST in the lower troposphere (ozone concentration at 0-50 ppbv), the upper
261 troposphere (ozone concentration at 50-150 ppbv) and the stratosphere (ozone concentration
262 at >150 ppbv).

263 Figure 2a-e shows the overall tropospheric ozone comparisons between independent
264 ozonesonde (Sonde-Observed) and Traj-Derived ozone in the entire study period (Figure 2a-c) and
265 each decade (Figures 2d). Overall, the Sonde-Observed and Traj-Derived ozone concentrations
266 agree well in the lower troposphere (Figure 2a), with a correlation coefficient (R) of 0.69 and a
267 root mean square (RMS) difference [square root of the mean of squared individual differences] of
268 7.5 ppbv, a low bias (0.7 ppbv) and RD (1.8%) [where RD is the relative difference $100 \times (\text{TOST}$
269 $\text{ozone} - \text{ozonesonde ozone}) / \text{ozonesonde ozone}$]. The linear fit for the entire study period shows a



270 slope of 0.99. In the upper troposphere (Figure 2b), the agreement between the Sonde-Observed
271 and Traj-Derived ozone concentration is moderately lower, with a linear fitting coefficient of 1.01
272 and RMS of 21.1 ppbv, and higher bias (2.9 ppbv) and RD (4.0%) than those in the lower
273 troposphere. This lower agreement in the upper troposphere owes to greater influence of
274 stratosphere-to-troposphere (STE) in the upper troposphere, where trajectories by the Lagrangian
275 dispersion model (such as HYSPLIT) show substantially increased deviations due to the strong
276 turbulence and convection (Stohl et al., 2002). The positive bias may imply that STE is slightly
277 overestimated in HYSPLIT, as the comparison between the Sonde-Observed and troposphere-only
278 Traj-Derived ozone concentrations shows a clear underestimation (with RD of -9% to -5%) in the
279 upper troposphere (Figure S2). In the stratosphere (Figure 2c), the overall agreement between the
280 Sonde-Observed and Traj-Derived ozone concentrations has a linear fitting coefficient of 0.97 and
281 an RMS of 416.9 ppbv. The small bias is of higher magnitude (11.1 ppbv) to that in the troposphere
282 but this is much smaller relative to stratospheric ozone concentrations; the RD is only 0.5%,
283 indicating higher reliability of Traj-Derived in the stratosphere.

284 This validation method compares ozonesonde station data with Traj-Derived ozone, i.e., the
285 ozone found by averaging trajectories that come from other stations, some of which will have
286 higher ozone, and some lower. The average difference results from an imbalance in the distribution
287 of meteorological trajectories, and this is confirmed by detailed analysis. For example, before the
288 1990s, fitting coefficients were smaller than 1 and Rs were smaller than 0.60 (Figure 2d) in the
289 lower troposphere, indicating a tendency to underestimate the Traj-Derived ozone in the lower
290 troposphere. After the 1990s, owing to the additional ozonesonde measurements provided by
291 SHADOZ in the tropics, the underestimation of Traj-Derived ozone in the lower troposphere is
292 greatly reduced and the linear fitting coefficient is very close to 1 (and Rs increased to > 0.71).
293 Similarly, with the additional ozonesonde measurements after the 1990s, the Rs in the upper



294 troposphere increased from < 0.50 to > 0.58 . In all decades, the agreement between Sonde-
295 Observed and Traj-Derived ozone in the stratosphere is the best, with Rs of ~ 0.97 and linear
296 coefficients of 0.99. The RD in each decade is small (-0.3% - 1.4%), indicating no systematic
297 underestimation or overestimation in the stratospheric Traj-Derived data. However, in the upper
298 troposphere, Traj-Derived ozone tends to be overestimated, with RD of 0.6–4.5%.

299 Figure 3 examines how the RD between the ozonesondes and Traj-Derived ozone values
300 varies with altitude, presenting the frequency distributions of RD across all stations, at every other
301 altitude level and in each decade. The distributions of RD show little skewness in every other
302 altitude and decade, indicating no systematic bias during the study period. The overall interquartile
303 ranges (25–75%) of RD are between -30 and 30% , with the lowest interquartile ranges of RD (-10
304 to 10%) in the stratosphere and middle and lower troposphere. Higher interquartile ranges of RD
305 appear in the 13–19 km altitude range, where the upper troposphere-lower stratosphere (UTLS)
306 region is located, and are due to the large vertical gradients of ozone concentrations in the UTLS
307 and the variability of the tropopause (Millan et al., 2023). The surface (boundary layer) ozone,
308 however, shows a positive bias of the median, in all decades, of up to 12% , suggesting that TOST,
309 which neglects ozone chemistry and deposition, often overestimates ozone concentration there.

310 Figure 4 exemplifies comparisons in vertical profiles between Sonde-Observed and Traj-
311 Derived ozone profiles at individual stations in different seasons. Four stations with sufficient data
312 coverage (>15 years) were selected from the Antarctic coastal region (Syowa), Europe
313 (Hohenpeissenberg), North America (Boulder), and East Asia (Beijing). The decadal mean (1990s
314 and 2000s) profiles in January and July are used to compare the performance of Traj-Derived
315 ozone profiles in boreal winter and summer. In general, the Traj-Derived profiles can capture the
316 vertical ozone variation in different seasons, with good correlation ($R > 0.99$) and high accuracy
317 (bias < 100 ppbv, RD $< 10\%$) in comparison to the independent ozonesonde profiles. The Syowa



318 comparison shows a larger bias, but much of this is due to the fact that in the 1990s this station
319 launched the Japanese KC-79 carbon-iodine sonde, while other stations in the Southern
320 Hemisphere launched ECC sondes; the Traj-Derived profiles would therefore be expected to be
321 10-20% higher in the troposphere and about 5% higher in the lower stratosphere (Smit and Kley,
322 1998). The excellent agreement in tropospheric ozone at Hohenpeissenberg is likely due to
323 frequent and dense European ozonesonde observations; similar cases also are seen at Uccle,
324 Payerne, and Praha. Larger discrepancies are shown near the planetary boundary layer (PBL) and
325 UTLS, as the simulated trajectories over these regions have more uncertainties (Stohl and Seibert,
326 1998; Sicard et al., 2019), and ozone chemistry and deposition are potentially important in the PBL
327 at time scales similar to that of the longer trajectories (four days).

328

329 **3.2 Comparisons with satellite data**

330 To compare with satellite data, we first validated the Traj-Derived ozone profiles against
331 ozonesonde measurements. The corresponding validation was conducted for the satellite data of
332 SAGE and MLS in the same period and location. The sets of ozonesonde, Traj-Derived and
333 satellite data were selected only when all three datasets were available in the same month, decade,
334 and gridpoint, so to ensure that both the Traj-Derived and satellite data could be independently
335 evaluated by the ozonesondes. Figures 5a-d show the vertical RD of the Traj-Derived and SAGE
336 ozone. Compared to SAGE, Traj-Derived ozone concentrations agree with the ozonesondes better
337 in the troposphere (<12 km), with the RD generally < 20%. Above 11 km, Traj-Derived and SAGE
338 ozone concentrations have comparable RD of 10-25% between 12-20 km, and less than 5% above
339 20 km. In the 12-20 km range, SAGE ozone agrees better with the ozonesondes, particularly in the
340 1980s.

341 Figures 5e-h compare the vertical RD of the Traj-Derived and MLS ozone values. The MLS



342 profiles are validated only above the altitude recommended (261 hPa, Livesey et al., 2022). In the
343 lowermost stratosphere, from 12-17 km, MLS shows comparable or better performance than Traj-
344 Derived ozone, while above 17 km the RD of MLS ozone is higher by 0.62-11.88% than that of
345 Traj-Derived ozone, particularly in boreal summer (JJA).

346 It is of course expected that TOST would outperform satellite instruments in measurements
347 below the tropopause, as satellite measurements are hampered by the large stratospheric ozone
348 burden that satellite instruments must look through, but these comparisons suggest that even above
349 15 km, where SAGE and MLS are considered most reliable (Wang et al., 2002; Kremser et al.,
350 2020; Livesey et al., 2022), TOST can provide comparable or better accuracy.

351 Figures 6 show time series of the vertical variation of monthly RD from 16-26 km between
352 Traj-Derived and SAGE ozone from 1985-2005, and between Traj-Derived and MLS ozone from
353 2005-2019. SAGE ozone data are reliable above 20 km (Kremser et al., 2020), having a mean RD
354 of about -10-10%, similar to that of Traj-Derived ozone. SAGE ozone concentrations are lower
355 than the Traj-Derived ozone by 5 to 10% between 16 and 20 km (Figure 6f), as Traj-Derived ozone
356 overestimates the ozonesondes by 9 to 15% (Figure 6e) while SAGE ozone underestimates the
357 ozonesondes by -7 to -1% (Figure 6d). Over the MLS period from 2005 to 2019, TOST ozone at
358 all altitudes between 16 and 26 km agrees with independent ozonesondes better than during the
359 SAGE period (Figures 6h and 6k vs. Figures 6b and 6e). Accordingly, the Traj-Derived ozone
360 concentrations show good agreement with MLS ozone above 22 km, but are lower than MLS
361 ozone below 20 km (Figures 6i and 6l), as MLS generally overestimates ozone concentrations
362 below 20 km (Figures 6g and 6j).

363 Figure S4 compares the RMSE of Traj-Derived and satellite ozone in different latitude zones
364 from 16-26 km. Compared to SAGE in the 1990s, the Traj-Derived ozone has comparable RMSEs
365 in the Northern Hemisphere, yet higher RMSEs in the Southern Hemisphere, due to the fewer



366 ozonesonde stations there. MLS ozone also shows lower RMSEs in the Southern Hemisphere, but
367 higher RMSEs in the Northern Hemisphere.

368 Figure S5 compares monthly average ozone mixing ratios of Traj-Derived ozone with
369 corresponding SAGE and MLS averages, above 16 km, in two seasons. The monthly average
370 values correlate very well, with $R = 0.94-0.98$, for both instruments and both seasons. Seasonally,
371 Traj-Derived ozone is slightly higher than either SAGE or MLS ozone in DJF (linear fitting
372 slope >1 ; RD between 1 and 3%), but markedly lower than MLS in JJA (linear fitting slope 0.91-
373 0.92; RD -8 to -9%).

374 Table S2 summarizes the evaluation of both Traj-Derived and satellite ozone against the
375 ozonesondes over 16-26 km. The Traj-Derived and SAGE ozone values show high correlation (R
376 $= 0.95$ or greater in all cases), and the Traj-Derived comparison shows RDs of -1% to +2% in the
377 1980s and 1990s, but only -0.3% to +0.4% in the 2000s and 2010s. By contrast, the SAGE
378 comparison shows RDs of -4% to +0.5%, while the MLS comparison shows RDs of -2% to +11%.
379

380 **3.3 Comparisons with aircraft observations**

381 We also compare TOST ozone with the IAGOS dataset, in the lower troposphere at 0-50 ppbv,
382 from 1994-2021 (Figure 7). Note that this comparison is between the full TOST (not Traj-derived)
383 and IAGOS datasets here. TOST ozone values are generally higher than IAGOS with a mean bias
384 of 2.2 ppbv and R of 0.49, but RDs (5.8%) and RMS (8.8 ppbv) are low. The linear fit has a slope
385 of 1.03. The two ozone datasets employ different measurement techniques and atmospheric
386 sampling (Petetin et al., 2018). Previous studies have reported that IAGOS ozone values are
387 systematically lower than ozonesonde values, typically by 5-10% in the free troposphere (Tilmes
388 et al., 2012; Zbinden et al., 2013; Staufer et al., 2013, 2014; Tanimoto et al., 2015; Tarasick et al.,
389 2019b). The comparisons in Figure 6 are consistent with these earlier estimates, as the RD (Figure



390 3d) indicates that IAGOS measurements average 6% lower than TOST, with only slight variation
391 (5-8%) when the comparison is made by decade. In the upper troposphere at 6-10 km, however,
392 the IAGOS measurements are on average 12% lower than TOST, with also slight variation (11-
393 12%) between decades (Figure S3).

394

395 **3.4 Improvements in the new version**

396 The improvements in TOST-v2 are attributed to the increased amount and improved quality of
397 ozonesonde data, as well as the improved trajectory simulation and ozone mapping. Because more
398 ozonesonde stations and more ozonesonde data have become available since the 1990s or 2000s
399 (Table S1), more ozone profiles were used in constructing TOST-2, leading to improved data
400 density. Table S3 summarizes the data coverage, the number of ozonesonde stations and
401 ozonesonde profiles used for TOST-v2 and TOST-v1. The data coverage is defined as the ratio of
402 the number of gridpoints with valid annual means to the total number of gridpoints in the
403 corresponding latitudinal zone. The number of ozonesonde stations, compared to Liu et al. (2013b),
404 increases in all latitudes by ~50%, and the total number of ozonesonde profiles used is doubled.
405 Data coverage increases as well, in all latitude bands, by 5-15% (Table S3) and in all altitudes by
406 a maximum of 10% (Figure S6).

407 In addition to the data density, the data quality was also improved in TOST-v2. Figure 8a-b
408 shows the distributions of ozone concentrations in TOST-v2 and TOST-v1 at the lowest level (0-1
409 km) for the 2000s. Over the Antarctic, gaps are observed only in the new TOST data. This is more
410 reasonable for the sea-level data because the altitude over the Antarctic is over 1 km (Figure S7a),
411 where ozone trajectories should not appear at 0-1 km. Therefore, the spatial distributions of ozone
412 are clearly improved with this topography correction in TOST-v2 compared to TOST-v1, which
413 could be attributed to the updated terrain file since HYSPLIT v5.0



414 (<https://www.arl.noaa.gov/hysplit/hysplit-model-updates/>). Over the eastern Pacific, marked with
415 an ellipse in Figure 8a, b, TOST-v1 shows higher ozone concentrations than TOST-v2 by 30%
416 (Figure 8c). Compared to the ozonesonde measurement at 0-1 km in the 2000s in these two regions
417 (Davis station for the Antarctic and Easter Island station for the eastern Pacific), TOST-v2 agrees
418 better with ozonesondes than TOST-v1, indicating better representation of ozone distributions
419 (Figure S7b).

420 With reference to spatial distributions at 19-20 km in the 2000s, Figure 8d-e shows that in the
421 Antarctic and the tropical eastern Pacific, TOST-v1 values show higher concentrations than TOST-
422 v2 (Figure 8f). Figure S7c compares ozone concentrations from ozonesonde, TOST-v2, and TOST-
423 v1 at 19-20 km in the 2000s at an Antarctic station (Syowa) and a tropical station (Bogota).
424 Compared to TOST-v1, TOST-v2 ozone values show a better agreement with the ozonesonde
425 measurement. The difference between TOST-v2 ozone and ozonesonde measurements is 10% and
426 29% in Syowa and Bogota stations, while in TOST-v1, ozone concentrations at these stations show
427 24% and 39% differences (Figure S7c).

428 In summary, TOST has been improved in TOST-v2 with higher spatial coverage, improved
429 description of ozone spatial distributions, and a better agreement with ozonesonde measurements
430 in both the troposphere and stratosphere.

431

432 **3.5 Uncertainty analysis**

433 As noted in Section 2.3, the ozone concentrations in each of the TOST gridpoints (or bins) in a
434 month are determined by the ozone concentrations along all the trajectories passing through that
435 gridpoint in that month. Therefore, an estimate of the random uncertainty of TOST may be
436 obtained from the standard error of the mean in each bin. Note that this may not be a true estimate
437 of the standard error, as some bins may contain more than one value from an individual trajectory,



438 depending on wind speed, and so these values are not independent and our standard error
439 calculation is biased low.

440 For convenience, given the large range of ozone concentrations between the stratosphere and
441 troposphere, we use the ratio of the standard error to the mean in that bin, SE/Mean, expressed
442 in %. The standard error is proportional to the variability of the ozone values in a bin (i.e. the
443 standard deviation) and inversely proportional to the square root of the number of data values.
444 Thus in general, the more trajectories passing a gridpoint, the more data points for that gridpoint
445 and the lower the standard error for that gridpoint. Figure 9 shows the SE/Mean and the number
446 of samples in January and July of the 2000s at 3-4 km and 19-20 km. Generally, the Southern
447 Hemisphere shows higher SE/Mean values ($> 10\%$) than the Northern Hemisphere ($< 6\%$), which
448 reflects the large number (> 100) of ozone soundings in the Northern Hemisphere, especially over
449 North America and Europe. However, near the equator, despite the higher sampling rate, the
450 SE/Mean still is as high as 15%. Compared to the stratospheric level (19-20 km), the tropospheric
451 level (3-4 km) shows an overall higher SE/Mean. SE/Mean varies less with season in the
452 stratosphere than in the troposphere. For example, at 3-4 km, the SE/Mean in January is generally
453 $< 7\%$ but becomes $> 10\%$ in July in the Northern Hemisphere, and vice versa in the Southern
454 Hemisphere. This is likely due to more vertical motion in the PBL (Stohl and Seibert, 1998; Sicard
455 et al., 2019) so that ozone in some bins comes from multiple altitude levels, as well as increased
456 photochemistry and biomass burning. Stratospheric intrusions to the lower troposphere are more
457 frequent in boreal spring and summer than in winter (Terao et al., 2008; Greenslade et al., 2017),
458 and can be responsible for much of the variability at 3-4 km (Tarasick et al., 2019a).

459 To quantify the uncertainties of TOST ozone in different altitudinal and latitudinal zones, and
460 in different seasons and decades, we calculated the Normalized Root Mean Squared Error
461 (NRMSE) of the monthly ozone mixing ratio between ozonesonde and Traj-Derived ozone over



1970-2021 (Figure 10). Among altitudes, the highest NRMSE values appear at 9-10 km and over the tropopause region, and the second highest NRMSE at the surface, while the lowest NRMSE values are in the lower to middle troposphere (3-6 km) and stratosphere (19-26 km), consistent with Figure 4. There is considerable variation in NRMSE with latitude; the NRMSEs in the southern high latitudes (90-60S) and the northern tropics (0-30N) are higher than in other latitudinal zones. This could reflect higher horizontal gradients of ozone (e.g. stations in or outside the ozone hole) in the southern high latitudes or biases between ECC sondes and other types (the Indian and Japanese sondes) in the northern tropics. By season, the NRMSE varies slightly with a lower value in March-April-May than in other seasons. After the 1990s, the NRMSEs are reduced markedly compared to the 1980s and 1970s, likely related to the improved data coverage in the later period. This overview provides caveats regarding where (surface and UTLS, the northern high latitudes and tropics) and when (before the 1990s) more caution is advised when using TOST.

4. Global ozone spatial-temporal variations observed from TOST

4.1 Ozone spatial variations in the troposphere and stratosphere

As a 3-dimensional ozone dataset, TOST can depict both horizontal and vertical ozone distributions, as well as long-term ozone timeseries. Figure 11 shows distributions of decadal mean TOST ozone at 3-4 km and 19-20 km in four seasons of the 2000s. At 3-4 km in the troposphere, ozone concentrations are higher over the continent in the Northern Hemisphere, especially in MAM and DJF (>50 ppbv), reflecting the ozone production from the photochemical reactions of anthropogenic and natural emissions. In addition, the continental outflow from the southern US (in MAM) and the biomass burning-produced ozone in southern Africa (in JJA and SON) are well captured and in agreement with satellite observations (Fishman et al., 1990; Ebojje et al., 2016). At 19-20 km in the stratosphere (Figure 11e-h), ozone concentrations are higher near the poles



486 than in the tropics, due to the impact of the Brewer–Dobson circulation. The North Pole has higher
487 ozone concentrations than the South Pole in DJF and MAM, and vice versa in JJA and SON,
488 reflecting the seasonality of the Brewer-Dobson circulation. Also at 19-20 km, the ozone
489 concentrations are lower over Asia in JJA (Figure 11f) than in other seasons, reflecting the
490 transport of ozone by Asian summer monsoon from the tropics (Gettelman et al., 2004; Bian et al.,
491 2020).

492 Although trajectory mapping fills in much of the spatial domain, large gaps can still be found,
493 particularly in the tropics, where ozone soundings are less dense. Since some applications require
494 a default ozone value at all gridpoints, a smoothed ozone dataset is also provided for the decadal
495 mean ozone in each month and the annual mean ozone, by fitting the maps at each level to a linear
496 combination of spherical functions (Liu et al., 2013b). As shown in Figures 11i-p, small-scale
497 variations and extreme values are reduced in the smoothed ozone fields, while broad patterns of
498 the ozone distribution are retained, making these smoothed maps valuable for qualitative
499 visualization of the spatial, seasonal, and decadal variations in ozone at different altitudes. They
500 should, however, be used for any kind of quantitative analysis with great caution, as these highly
501 interpolated data, where gaps exist in the unsmoothed TOST dataset, are necessarily far from any
502 original measurement and the degree to which they represent the true ozone value is doubtful. For
503 example, erroneous conclusions have been inferred from the smoothed TOST-v1 output over the
504 tropics, with very limited observations before 1998, where the smoothed data were mostly
505 interpolated from higher latitudes (Chipperfield et al., 2022). In addition, smoothing, as noted,
506 removes small-scale variations and extreme values, and does so whether they are real or not. The
507 smoothed dataset has not been quantitatively evaluated in any way.

508 Figure 12a-d shows the latitude-altitude distribution of TOST ozone in each season averaged
509 over 1970-2021. The steep changes in ozone concentration from <100 to >500 ppbv in the vicinity



510 of the tropopause (the black lines in Figure 12a-d, calculated from the NCEP/NCAR reanalysis)
511 are well captured. Due to the Brewer-Dobson circulation, ozone concentrations above the
512 tropopause increase with latitude from the tropics to the poles, which is also well reflected in the
513 latitude-altitude distribution. TOST ozone concentrations are higher in spring (600-800 ppbv) than
514 in the other seasons (< 500 ppbv) over northern midlatitudes (45-60°N) at about 12-13 km, which
515 reflects the stronger Brewer-Dobson circulation in spring (Holton et al., 1995). Figure 12e shows
516 the monthly mean TOST ozone time series from 1970 to 2021, averaged over 30-70°N at each
517 level. Clear seasonal cycles are well captured every year.

518

519 **4.2 Long-term trend in stratospheric ozone**

520 One of the advantages of TOST is its long-term coverage, which enables investigation of variations
521 in ozone back to the 1970s. One application is to study stratospheric ozone changes, as it is
522 important to assess stratospheric ozone recovery (or lack thereof) since the implementation of the
523 Montreal Protocol and its amendments (Fang et al., 2019). While this is commonly done with
524 individual ozonesonde time series, it is challenging to assess how well individual long-term station
525 changes represent regional or global variations. Combining data records from sparse and widely
526 separated ozonesonde sites involves implicit assumptions about their representativeness. With
527 meteorological trajectory mapping, each original ozonesonde measurement is assigned a trajectory
528 which describes its representativeness, and the TOST averages are therefore weighted according
529 to the representativeness of each measurement. While this is subject to trajectory errors and the
530 fact that coverage is incomplete (Table S3), unless trajectory errors are non-random, it should
531 produce a better result than simple averaging of sonde station data by geographic region.

532 Figure 13 shows the area-weighted annual averages of ozone concentrations at 21-22 km and
533 24-25 km from 1970 to 2021; averages were taken over all gridpoints from 30°-70°N with



534 available data throughout all years (~ 70% of gridpoints). The 3-year running means are also shown
535 with the time series. The ozone time series at both levels captures the ozone depletion in the early
536 1990s from the effects of the 1991 eruption of Mt. Pinatubo (McCormick et al., 1995; Tang et al.,
537 2013; Dhomse, et al., 2015) and the recovery in the latter part of the 1990s. In addition, these
538 updated TOST time series show that stratospheric ozone since 2000 changed little, despite the
539 decline in stratospheric chlorine since then. From Figure 14, there is an insignificant trend in the
540 ozone concentrations at 21-22 km (by 0.6 ppbv/year) and 24-25 km (by -1.9 ppbv/year) from 1998
541 to 2021, indicating little change of stratospheric ozone, despite the fact that 25 years have passed
542 since peak stratospheric chlorine. Using long-term satellite data, Bognar et al. (2022) also find
543 stratospheric ozone largely unchanged in the last two decades, which, they suggest, is related to
544 asymmetries and long-term variability in the Brewer-Dobson circulation. Such observations of the
545 variation in stratospheric ozone are essential to verifying the expected stratospheric ozone recovery
546 under the Montreal Protocol.

547

548 **5. Conclusions**

549 An improved TOST dataset has been generated from 1970 to 2021 based on the updated
550 ozonesonde profiles at 141 ozonesonde stations from WOUDC, SHADOZ, HEGIFTOM and
551 NDACC. The updated TOST was derived by combining the 4-day forward and backward
552 trajectories from each ozonesonde profile, which were driven by the most state-of-the-art
553 HYSPLIT model (v5.2) and NCEP reanalysis data (NNRP-1). The monthly mean ozone, decadal
554 mean ozone in each month (January to December from the 1970s-2010s) and annual mean ozone
555 (1970-2021) are provided in 3-dimensional grids of $5^{\circ} \times 5^{\circ} \times 1$ km (latitude, longitude, and
556 altitude). Spatially-smoothed maps are also provided by decadal mean in each month and annual
557 mean for qualitative visualization, model initialization, and other applications with caution. For



558 user convenience, the TOST data in coordinates from the sea level and from the surface level are
559 both generated, and separate ozone climatology datasets are generated based on trajectories from
560 ozonesonde in both the troposphere and stratosphere, trajectories from ozonesonde only in the
561 troposphere and trajectories from ozonesonde only in the stratosphere. Statistics (standard error,
562 number of samples) are also provided.

563 Comprehensive validation of TOST-v2 was conducted. At all the ozonesonde stations used,
564 trajectory-derived ozone profiles without the input of the station itself were compared with the
565 corresponding ozonesonde profiles at the stations. The overall comparison between the
566 ozonesonde and trajectory-derived ozone shows good agreement in both the troposphere ($R = 0.56$ -
567 0.69 , $RD = 2$ - 4%) and stratosphere ($R = 0.97$, $RD = 0.5\%$) in each decade and in all decades' mean
568 (Figure 2). The frequency distribution of RD at different altitudes shows interquartile ranges of
569 RD between -30 to 30% , with the lowest interquartile ranges of RD (-10 to 10%) in the stratosphere
570 and lower troposphere, and no systematic bias except in the surface layer (Figure 3). The patterns
571 of ozone profiles at individual stations are also well captured and quantified, with $R > 0.76$ and
572 RD of 2 - 8% (Figure 4). Larger discrepancies are shown near the PBL and UTLS, especially for
573 coastal stations where the trajectory-derived ozone may be biased by trajectories from the
574 continent (Tarasick et al., 2010).

575 The comparison between TOST and satellite data, i.e., SAGE in the 1980s and 1990s, and
576 MLS in the 2000s and 2010s, illustrates that TOST data have comparable accuracy with the
577 satellite data in the stratosphere, while in the troposphere TOST is markedly superior (Figure 5).
578 In different latitude zones and decades, TOST performs comparably with SAGE and MLS data as
579 well (Figure 6). TOST-v2 was also directly compared to MOZAIC-IAGOS ozone profiles over the
580 period 1994-2021 from the surface to 5 km. Despite the systematic difference between MOZAIC-
581 IAGOS and ozonesonde measurements, the two ozone datasets agree well in each decade and in



582 all decades' mean for lower troposphere ($RD = 5-8\%$, Figure 7) and upper troposphere ($RD = 11-$
583 12% , Figure S3).

584 Compared to the previous version of TOST (TOST-v1, Liu et al. 2013a and b), this new TOST,
585 TOST-v2, is improved in two major aspects. Firstly, the record is extended to 2021 and data
586 coverage is increased by as much as 15% , as more ozone profiles and 43 additional ozonesonde
587 stations are used in constructing the new version of TOST. Secondly, the spatial distribution of
588 ozone has better agreement with ozonesonde measurements in both the troposphere and
589 stratosphere over regions of Antarctica and the eastern Pacific, with RD decreased by $> 50\%$.
590 The uncertainties of TOST are largely dependent on the availability of ozonesonde data. Higher
591 uncertainties are found before the 1990s, as global coverage is sparse in the tropics before
592 SHADOZ. Higher uncertainties also appear at southern high latitudes and in the northern tropics,
593 with NRMSE there being $35-78\%$ higher than at northern midlatitudes (Figure 10), likely because
594 of greater ozone variability there, although biases between ozonesonde types may also contribute.
595 TOST data at the PBL and UTLS have higher standard error and twice the NRMSE compared to
596 other altitude levels; the former is due to more small-scale processes in the PBL while the latter is
597 related to the large ozone gradient and the dynamic variation of the tropopause.

598 TOST can capture global ozone distributions in the troposphere and stratosphere (Figures 11
599 and 12), showing horizontal and vertical variations, the continental outflow, and the gradient of
600 ozone concentration near the tropopause. TOST can also reflect the seasonal variations in ozone
601 concentrations near the vicinity of the tropopause. The time series of the updated TOST shows the
602 stagnant recovery but overall insignificant change of stratospheric ozone after 1998 (Figure 13),
603 which agrees well with studies using satellite-based and model-based ozone datasets (Bognar et
604 al., 2022).

605 It is anticipated that this updated and improved TOST dataset can benefit future studies, owing



606 to its long record, global coverage, and high vertical resolution. We expect that it will be a useful
607 dataset for trend studies, especially in the free troposphere, and also in the stratosphere, given the
608 excellent long-term stability of the global ozonesonde network (Stauffer et al., 2022). We caution,
609 however, that users should keep in mind the assumptions and limitations of the data product as
610 described here.

611

612 **Author contribution**

613 J. L. and D.T. conceptualized and designed this study. Z.Z. performed data process, analysis, and
614 composed the first draft. All the coauthors contributed substantially to this study in making
615 ozonesonde measurements, processing, calibrating, and archiving the ozonesonde data, and
616 providing constructive and valuable suggestions to and comments on the manuscript. All the co-
617 authors approved the submission of this paper.

618

619 **Code and data availability**

620 The ozonesonde data used in this study can be obtained from the WOUDC
621 (https://woudc.org/archive/Archive-NewFormat/OzoneSonde_1.0_1/), SHADOZ
622 (<https://doi.org/10.57721/SHADOZ-V06>) and HEGIFTOM ([https://hegiftom.meteo.be/datasets](https://hegiftom.meteo.be/datasets/ozonesondes)
623 [/ozonesondes](https://hegiftom.meteo.be/datasets/ozonesondes)). The trajectory model HYSPLIT (Version 5.2) is from the NOAA Air Resources
624 Laboratory (<http://www.arl.noaa.gov/ready.html>), driven by the NCEP/NCAR reanalysis data
625 from the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, at [https://www.ready.noaa.gov/data/](https://www.ready.noaa.gov/data/archives/reanalysis/)
626 [archives/reanalysis/](https://www.ready.noaa.gov/data/archives/reanalysis/). The aircraft data can be accessed from IAGOS network
627 (<https://www.iagos.org/>). The two satellite data for comparison, the SAGE II (Version 7.0) and
628 the MLS (Version 5.0), are obtained from <https://sage.nasa.gov/missions/about-sage-ii/> and
629 https://disc.gsfc.nasa.gov/datasets/ML2O3_005/summary?keywords=ML2O3_005, respectively.



630 We are in the process of making the TOST available at the WOUDC website. TOST data
631 currently are available on request from the authors.

632

633 **Competing interests**

634 The authors declare that they have no conflict of interest.

635

636 **Special issue statement**

637 This article is part of the special issue “Tropospheric Ozone Assessment Report Phase II (TOAR-
638 II) Community Special Issue (ACP/AMT/BG/GMD inter-journal SI)”.

639

640 **Acknowledgements**

641 We thank many for their dedication to WOUDC, SHADOZ, and HEGIFTOM, making ozonesonde
642 data accessible. We also thank SAGE II and MLS team for their ozone data for comparison. We
643 acknowledge the HYSPLIT team for the trajectory model. Z. Z. and J. L. acknowledge the financial
644 support from Natural Science and Engineering Council of Canada (Grant No. RGPIN-2020-
645 05163); J. B. and J. Z. from the National Natural Science Foundation of China (Grant No.
646 42293321).

647

648 **References**

649 Ancellet, G., Godin-Beekmann, S., Smit, H. G., Stauffer, R. M., Van Malderen, R., Bodichon, R.,
650 and Pazmiño, A.: Homogenization of the Observatoire de Haute Provence electrochemical
651 concentration cell (ECC) ozonesonde data record: comparison with lidar and satellite
652 observations, *Atmospheric Measurement Techniques*, 15, 3105-3120, 2022.
653 Badia, A., Iglesias-Suarez, F., Fernandez, R. P., Cuevas, C. A., Kinnison, D. E., Lamarque, J. F.,
654 Griffiths, P. T., Tarasick, D. W., Liu, J., and Saiz-Lopez, A.: The role of natural halogens in



655 global tropospheric ozone chemistry and budget under different 21st century climate
656 scenarios, *Journal of Geophysical Research: Atmospheres*, 126, e2021JD034859, 2021.

657 Bhartia, P. K.: OMI Algorithm Theoretical Basis Document Volume II, OMI Ozone, 2002.

658 Bian, J., Gettelman, A., Chen, H., and Pan, L. L.: Validation of satellite ozone profile retrievals
659 using Beijing ozonesonde data, *Journal of Geophysical Research: Atmospheres*, 112, 2007.

660 Bian, J., Li, D., Bai, Z., Li, Q., Lyu, D., and Zhou, X.: Transport of Asian surface pollutants to the
661 global stratosphere from the Tibetan Plateau region during the Asian summer monsoon,
662 *National Science Review*, 7, 516-533, 2020.

663 Bodeker, G. E., Nitzbon, J., Tradowsky, J. S., Kremser, S., Schwertheim, A., and Lewis, J.: A global
664 total column ozone climate data record, *Earth System Science Data*, 13, 3885-3906, 2021.

665 Bogner, K., Tegtmeier, S., Bourassa, A., Roth, C., Warnock, T., Zawada, D., and Degenstein, D.:
666 Stratospheric ozone trends for 1984–2021 in the SAGE II–OSIRIS–SAGE III/ISS composite
667 dataset, *Atmospheric Chemistry and Physics*, 22, 9553-9569, 2022.

668 Chen, X., Liu, Y., Lai, A., Han, S., Fan, Q., Wang, X., Ling, Z., Huang, F., and Fan, S.: Factors
669 dominating 3-dimensional ozone distribution during high tropospheric ozone period,
670 *Environmental Pollution*, 232, 55-64, 2018.

671 Chipperfield, M. P., Chrysanthou, A., Damadeo, R., Dameris, M., Dhomse, S. S., Fioletov, V., Frith,
672 S. M., Godin-Beekmann, S., Hassler, B., and Liu, J.: Comment on “Observation of large and
673 all-season ozone losses over the tropics” [AIP Adv. 12, 075006 (2022)], *AIP Advances*, 12,
674 2022.

675 Cunnold, D., Wang, H., Chu, W., and Froidevaux, L.: Comparisons between Stratospheric Aerosol
676 and Gas Experiment II and microwave limb sounder ozone measurements and aliasing of
677 SAGE II ozone trends in the lower stratosphere, *Journal of Geophysical Research:*
678 *Atmospheres*, 101, 10061-10075, 1996.

679 Dhomse, SS, MP Chipperfield, W Feng, R Hossaini, GW Mann, and ML Santee, Revisiting the
680 hemispheric asymmetry in midlatitude ozone changes following the Mount Pinatubo eruption:
681 A 3-D model study. *Geophysical Research Letters*, 42, 3038–3047, 2015.

682 Draxler, R. R. and Hess, G.: An overview of the HYSPLIT_4 modelling system for trajectories,
683 *Australian meteorological magazine*, 47, 295-308, 1998.

684 Ebojie, F., Burrows, J., Gebhardt, C., Ladstätter-Weissenmayer, A., Von Savigny, C., Rozanov, A.,
685 Weber, M., and Bovensmann, H.: Global tropospheric ozone variations from 2003 to 2011 as



- 686 seen by SCIAMACHY, *Atmospheric Chemistry and Physics*, 16, 417-436, 2016.
- 687 Engström, A. and Magnusson, L.: Estimating trajectory uncertainties due to flow dependent errors
688 in the atmospheric analysis, *Atmospheric Chemistry and Physics*, 9, 8857-8867, 2009.
- 689 Eyring, V., Cionni, I., Bodeker, G. E., Charlton-Perez, A. J., Kinnison, D. E., Scinocca, J. F., Waugh,
690 D. W., Akiyoshi, H., Bekki, S., and Chipperfield, M. P.: Multi-model assessment of
691 stratospheric ozone return dates and ozone recovery in CCMVal-2 models, *Atmospheric*
692 *Chemistry and Physics*, 10, 9451-9472, 2010.
- 693 Fang, X., Pyle, J. A., Chipperfield, M. P., Daniel, J. S., Park, S., and Prinn, R. G.: Challenges for
694 the recovery of the ozone layer, *Nature Geoscience*, 12, 592-596, 2019.
- 695 Fioletov, V., Tarasick, D., and Petropavlovskikh, I.: Estimating ozone variability and instrument
696 uncertainties from SBUV ($/2$), ozonesonde, Umkehr, and SAGE II measurements: Short-term
697 variations, *Journal of Geophysical Research: Atmospheres*, 111, 2006.
- 698 Fishman, J., Watson, C. E., Larsen, J. C., and Logan, J. A.: Distribution of tropospheric ozone
699 determined from satellite data, *Journal of Geophysical Research: Atmospheres*, 95, 3599-
700 3617, 1990.
- 701 Gaudel, A., Cooper, O. R., Ancellet, G., Barret, B., Boynard, A., Burrows, J. P., Clerbaux, C.,
702 Coheur, P.-F., Cuesta, J., and Cuevas, E.: Tropospheric Ozone Assessment Report: Present-
703 day distribution and trends of tropospheric ozone relevant to climate and global atmospheric
704 chemistry model evaluation, *Elementa: Science of the Anthropocene*, 6, 2018.
- 705 Gettelman, A., Kinnison, D. E., Dunkerton, T. J., and Brasseur, G. P.: Impact of monsoon
706 circulations on the upper troposphere and lower stratosphere, *Journal of Geophysical*
707 *Research: Atmospheres*, 109, 2004.
- 708 Greenslade, J. W., Alexander, S. P., Schofield, R., Fisher, J. A., and Klekociuk, A. K.: Stratospheric
709 ozone intrusion events and their impacts on tropospheric ozone in the Southern Hemisphere,
710 *Atmospheric Chemistry and Physics*, 17, 10269-10290, 2017.
- 711 Griffiths, P. T., Murray, L. T., Zeng, G., Shin, Y. M., Abraham, N. L., Archibald, A. T., Deushi, M.,
712 Emmons, L. K., Galbally, I. E., Hassler, B. L.W. Horowitz, J. Keeble, J. Liu, O. Moeini, V.
713 Naik, F.M. O'Connor, D. Tarasick, S. Tilmes, S.T. Turnock, O. Wild, P.J. Young and P. Zanis:
714 Tropospheric ozone in CMIP6 simulations, *Atmospheric Chemistry and Physics*, 21, 4187-
715 4218, 2021.
- 716 Gulev, S. K., Thorne, P. W., Ahn, J., Dentener, F. J., Domingues, C. M., Gerland, S., Gong, D.,



- 717 Kaufman, D. S., Nnamchi, H. C., and Quaas, J.: Changing state of the climate system, 2021.
- 718 Hassler, B., Bodeker, G., and Dameris, M.: A new global database of trace gases and aerosols from
- 719 multiple sources of high vertical resolution measurements, *Atmospheric Chemistry and*
- 720 *Physics*, 8, 5403-5421, 2008.
- 721 Hassler, B., Kremser, S., Bodeker, G. E., Lewis, J., Nesbit, K., Davis, S. M., Chipperfield, M. P.,
- 722 Dhomse, S. S., and Dameris, M.: An updated version of a gap-free monthly mean zonal mean
- 723 ozone database, *Earth System Science Data*, 10, 1473-1490, 2018.
- 724 Holton, J. R., Haynes, P. H., McIntyre, M. E., Douglass, A. R., Rood, R. B., and Pfister, L.:
- 725 Stratosphere-troposphere exchange, *Reviews of Geophysics*, 33, 403-439, 1995.
- 726 Jacob, D. J.: *Introduction to atmospheric chemistry*, Princeton University Press, 1999.
- 727 Kalney, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S.,
- 728 White, G., Woollen, J., Zhu, Y., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo,
- 729 K. C., Ropelewski, C., Wang, J., Leetmaa, A., Reynolds, R., Jenne, R., and Joseph, D.: The
- 730 NCEP/NCAR 40-Year Reanalysis Project, *Bulletin of the American Meteorological Society*,
- 731 77, 437-472, 1996.
- 732 Kent, G., Winker, D., Osborn, M., and Skeens, K.: A model for the separation of cloud and aerosol
- 733 in SAGE II occultation data, *Journal of Geophysical Research: Atmospheres*, 98, 20725-
- 734 20735, 1993.
- 735 Kremser, S., Thomason, L. W., and Bird, L. J.: Simplified SAGE II ozone data usage rules, *Earth*
- 736 *System Science Data*, 12, 1419-1435, 2020.
- 737 Liu, G., Tarasick, D. W., Fioletov, V. E., Sioris, C. E., and Rochon, Y. J.: Ozone correlation lengths
- 738 and measurement uncertainties from analysis of historical ozonesonde data in North America
- 739 and Europe, *Journal of Geophysical Research: Atmospheres*, 114, 2009.
- 740 Liu, G., Liu, J., Tarasick, D., Fioletov, V., Jin, J., Moeini, O., Liu, X., Sioris, C., and Osman, M.:
- 741 A global tropospheric ozone climatology from trajectory-mapped ozone soundings,
- 742 *Atmospheric Chemistry and Physics*, 13, 10659-10675, 2013a.
- 743 Liu, J., Tarasick, D., Fioletov, V., McLinden, C., Zhao, T., Gong, S., Sioris, C., Jin, J., Liu, G., and
- 744 Moeini, O.: A global ozone climatology from ozone soundings via trajectory mapping: a
- 745 stratospheric perspective, *Atmospheric Chemistry and Physics*, 13, 11441-11464, 2013b.
- 746 Liu, X., Bhartia, P., Chance, K., Spurr, R., and Kurosu, T.: Ozone profile retrievals from the Ozone
- 747 Monitoring Instrument, *Atmospheric Chemistry and Physics*, 10, 2521-2537, 2010.



- 748 Livesey, N., Read, W., Wagner, P., Froidevaux, L., Santee, M., Schwartz, M., Lambert, A., Millán
749 Valle, L., Pumphrey, H., and Manney, G.: Earth Observing System (EOS) Aura Microwave
750 Limb Sounder (MLS) version 5.0 x level 2 and 3 data quality and description document
751 Version 5.0–1.1 a (Tech. Rep.), Jet Propulsion Laboratory, California Institute of Technology.
752 Retrieved from https://mls.jpl.nasa.gov/data/v5-0_data_quality_document.pdf, 2022.
- 753 McCormick, J. P., L. W. Thomason, and C. R. Trepte (1995), Atmospheric effects of the Mt.
754 Pinatubo eruption, *Nature*, 373, 399–404.
- 755 McPeters, R. D. and Labow, G. J.: Climatology 2011: An MLS and sonde derived ozone
756 climatology for satellite retrieval algorithms, *Journal of Geophysical Research: Atmospheres*,
757 117, 2012.
- 758 McPeters, R. D., Labow, G. J., and Logan, J. A.: Ozone climatological profiles for satellite retrieval
759 algorithms, *Journal of Geophysical Research: Atmospheres*, 112, 2007.
- 760 Millan, L. F., Manney, G. L., Boenisch, H., Hegglin, M. I., Hoor, P., Kunkel, D., Leblanc, T.,
761 Petropavlovskikh, I., Walker, K., and Wargan, K.: A Multi-Parameter Dynamical Diagnostics
762 for Upper Tropospheric and Lower Stratospheric Studies, *EGUsphere*, 1–41, 2023.
- 763 Moeini, O., Tarasick, D. W., McElroy, C. T., Liu, J., Osman, M. K., Thompson, A. M., Parrington,
764 M., Palmer, P. I., Johnson, B., and Oltmans, S. J.: Estimating wildfire-generated ozone over
765 North America using ozonesonde profiles and a differential back trajectory technique,
766 *Atmospheric Environment: X*, 7, 100078, 2020.
- 767 Nédélec, P., Blot, R., Boulanger, D., Athier, G., Cousin, J.-M., Gautron, B., Petzold, A., Volz-
768 Thomas, A., and Thouret, V.: Instrumentation on commercial aircraft for monitoring the
769 atmospheric composition on a global scale: the IAGOS system, technical overview of ozone
770 and carbon monoxide measurements, *Tellus B: Chemical and Physical Meteorology*, 67,
771 27791, 2015.
- 772 Nowack, P. J., Luke Abraham, N., Maycock, A. C., Braesicke, P., Gregory, J. M., Joshi, M. M.,
773 Osprey, A., and Pyle, J. A.: A large ozone-circulation feedback and its implications for global
774 warming assessments, *Nature Climate Change*, 5, 41–45, 2015.
- 775 Petetin, H., Sauvage, B., Smit, H. G., Gheusi, F., Lohou, F., Blot, R., Clark, H., Athier, G.,
776 Boulanger, D., and Cousin, J.-M.: A climatological view of the vertical stratification of RH,
777 O₃ and CO within the PBL and at the interface with free troposphere as seen by IAGOS
778 aircraft and ozonesondes at northern mid-latitudes over 1994–2016, *Atmospheric Chemistry*



- 779 and Physics, 18, 9561-9581, 2018.
- 780 Petzold, A., Thouret, V., Gerbig, C., Zahn, A., Brenninkmeijer, C. A., Gallagher, M., Hermann, M.,
781 Pontaud, M., Ziereis, H., and Boulanger, D.: Global-scale atmosphere monitoring by in-
782 service aircraft—current achievements and future prospects of the European Research
783 Infrastructure IAGOS, *Tellus B: Chemical and Physical Meteorology*, 67, 28452, 2015.
- 784 Polvani, L. M., Wang, L., Aquila, V., and Waugh, D. W.: The impact of ozone-depleting substances
785 on tropical upwelling, as revealed by the absence of lower-stratospheric cooling since the late
786 1990s, *Journal of Climate*, 30, 2523-2534, 2017.
- 787 Sicard, M., Granados-Muñoz, M. J., Alados-Arboledas, L., Barragán, R., Bedoya-Velásquez, A.
788 E., Benavent-Oltra, J. A., Bortoli, D., Comerón, A., Córdoba-Jabonero, C., and Costa, M. J.:
789 Ground/space, passive/active remote sensing observations coupled with particle dispersion
790 modelling to understand the inter-continental transport of wildfire smoke plumes, *Remote
791 Sensing of Environment*, 232, 111294, 2019.
- 792 Skeie, R. B., Myhre, G., Hodnebrog, Ø., Cameron-Smith, P. J., Deushi, M., Hegglin, M. I.,
793 Horowitz, L. W., Kramer, R. J., Michou, M., and Mills, M. J.: Historical total ozone radiative
794 forcing derived from CMIP6 simulations, *Npj Climate and Atmospheric Science*, 3, 32, 2020.
- 795 Smit, H. G. and Kley, D.: JOSIE: The 1996 WMO International intercomparison of ozonesondes
796 under quasi flight conditions in the environmental simulation chamber at Jülich, *Proceedings
797 of the XVIII Quadrennial Ozone Symposium*, 971-974, 1998.
- 798 Smit, H. and Thompson, A.: Ozone-sonde Measurement Principles and Best Operational Practices:
799 ASOPOS 2.0 (Assessment of Standard Operating Procedures for Ozone-sondes), WMO,
800 World Meteorological Organization, GAW Report, 173, 2021.
- 801 Stauffer, J., Staehelin, J., Stübi, R., Peter, T., Tummon, F., and Thouret, V.: Trajectory matching of
802 ozonesondes and MOZAIC measurements in the UTLS — Part 1: Method description
803 and application at Payerne, Switzerland, *Atmospheric Measurement Techniques*, 6, 3393-
804 3406, 10.5194/amt-6-3393-2013, 2013.
- 805 Stauffer, J., Stähelin, J., Stübi, R., Peter, T., Tummon, F., and Thouret, V.: Trajectory matching of
806 ozonesondes and MOZAIC measurements in the UTLS—Part 2: Application to the global
807 ozonesonde network, *Atmospheric Measurement Techniques*, 7, 241-266, 2014.
- 808 Stauffer, R. M., Thompson, A. M., Kollonige, D. E., Tarasick, D. W., Van Malderen, R., Smit, H.
809 G., Vömel, H., Morris, G. A., Johnson, B. J., and Cullis, P. D.: An examination of the recent



- 810 stability of ozonesonde global network data, *Earth and Space Science*, 9, e2022EA002459,
811 2022.
- 812 Stauffer, R. M., Thompson, A. M., Kollonige, D. E., Witte, J. C., Tarasick, D. W., Davies, J., Vömel,
813 H., Morris, G. A., Van Malderen, R., and Johnson, B. J.: A post-2013 dropoff in total ozone
814 at a third of global ozonesonde stations: Electrochemical concentration cell instrument
815 artifacts? *Geophysical Research Letters*, 47, e2019GL086791, 2020.
- 816 Stein, A., Draxler, R. R., Rolph, G. D., Stunder, B. J., Cohen, M., and Ngan, F.: NOAA's HYSPLIT
817 atmospheric transport and dispersion modeling system, *Bulletin of the American
818 Meteorological Society*, 96, 2059-2077, 2015.
- 819 Steinbrecht, W., Froidevaux, L., Fuller, R., Wang, R., Anderson, J., Roth, C., Bourassa, A.,
820 Degenstein, D., Damadeo, R., and Zawodny, J.: An update on ozone profile trends for the
821 period 2000 to 2016, *Atmospheric Chemistry and Physics*, 17, 10675-10690, 2017.
- 822 Sterling, C. W., Johnson, B. J., Oltmans, S. J., Smit, H. G., Jordan, A. F., Cullis, P. D., Hall, E. G.,
823 Thompson, A. M., and Witte, J. C.: Homogenizing and estimating the uncertainty in NOAA's
824 long-term vertical ozone profile records measured with the electrochemical concentration cell
825 ozonesonde, *Atmospheric Measurement Techniques*, 11, 3661-3687, 2018.
- 826 Stohl, A.: Computation, accuracy and applications of trajectories—A review and bibliography,
827 *Atmospheric Environment*, 32, 947-966, 1998.
- 828 Stohl, A., Eckhardt, S., Forster, C., James, P., Spichtinger, N., and Seibert, P.: A replacement for
829 simple back trajectory calculations in the interpretation of atmospheric trace substance
830 measurements, *Atmospheric Environment*, 36, 4635-4648, 2002.
- 831 Stohl, A., James, P., Forster, C., Spichtinger, N., Marenco, A., Thouret, V., and Smit, H. G.: An
832 extension of Measurement of Ozone and Water Vapour by Airbus In-service Aircraft
833 (MOZAIC) ozone climatologies using trajectory statistics, *Journal of Geophysical Research:
834 Atmospheres*, 106, 27757-27768, 2001.
- 835 Stohl, A. and Seibert, P.: Accuracy of trajectories as determined from the conservation of
836 meteorological tracers, *Quarterly Journal of the Royal Meteorological Society*, 124, 1465-
837 1484, 1998.
- 838 Tang, Q., P. G. Hess, B. Brown-Steiner, and D. E. Kinnison, Tropospheric ozone decrease due to
839 the Mount Pinatubo eruption: Reduced stratospheric influx, *Geophysical Research
840 Letters*, 40, 5553–5558, 2013.



- 841 Tanimoto, H., Zbinden, R. M., Thouret, V., and Nédélec, P.: Consistency of tropospheric ozone
842 observations made by different platforms and techniques in the global databases, *Tellus B:*
843 *Chemical and Physical Meteorology*, 67, 27073, 2015.
- 844 Tarasick, D., Carey-Smith, T., Hocking, W., Moeini, O., He, H., Liu, J., Osman, M., Thompson,
845 A., Johnson, B., and Oltmans, S.: Quantifying stratosphere-troposphere transport of ozone
846 using balloon-borne ozonesondes, radar windprofilers and trajectory models, *Atmospheric*
847 *Environment*, 198, 496-509, 2019a.
- 848 Tarasick, D., Jin, J., Fioletov, V., Liu, G., Thompson, A., Oltmans, S., Liu, J., Sioris, C., Liu, X.,
849 and Cooper, O.: High-resolution tropospheric ozone fields for INTEX and ARCTAS from
850 IONS ozonesondes, *Journal of Geophysical Research: Atmospheres*, 115, 2010.
- 851 Tarasick, D.W., I. Galbally, O.R. Cooper, M.G. Schultz, G. Ancellet, T. LeBlanc, T.J. Wallington,
852 J. Ziemke, X. Liu, M. Steinbacher, J. Stähelin, C. Vigouroux, J. Hannigan, O. García, G. Foret,
853 P. Zanis, E. Weatherhead, I. Petropavlovskikh, H. Worden, J.L. Neu, M. Osman, J. Liu, M.
854 Lin, M. Granados-Muñoz, A.M. Thompson, S.J. Oltmans, J. Cuesta, G. Dufour, V. Thouret,
855 B. Hassler, A.M. Thompson and T. Trickl, TOAR- Observations: Tropospheric ozone from
856 1877 to 2016, observed levels, trends and uncertainties, *Elementa: Science of the*
857 *Anthropocene*, 7(1), p.39, 2019b.
- 858 Tarasick, D.W., J. Davies, H.G.J. Smit and S.J. Oltmans, A re-evaluated Canadian ozonesonde
859 record: measurements of the vertical distribution of ozone over Canada from 1966 to 2013,
860 *Atmospheric Measurement Techniques*, 9, 195-214, 2016.
- 861 Tarasick, D.W., H.G.J. Smit, A.M. Thompson G.A. Morris, J.C. Witte, J. Davies, T. Nakano, R.
862 van Malderen, R.M. Stauffer, T. Deshler, B.J. Johnson, R. Stübi, S.J. Oltmans and H. Vömel,
863 Improving ECC ozonesonde data quality: Assessment of current methods and outstanding
864 issues. *Earth and Space Science*, 8, e2019EA000914, 2021.
- 865 Terao, Y., Logan, J. A., Douglass, A. R., and Stolarski, R. S.: Contribution of stratospheric ozone
866 to the interannual variability of tropospheric ozone in the northern extratropics, *Journal of*
867 *Geophysical Research: Atmospheres*, 113, 2008.
- 868 Thompson, A. M., Witte, J. C., Sterling, C., Jordan, A., Johnson, B. J., Oltmans, S. J., Fujiwara,
869 M., Vömel, H., Allaart, M., and Peters, A.: First reprocessing of Southern Hemisphere
870 Additional Ozonesondes (SHADOZ) ozone profiles (1998–2016): 2. Comparisons with
871 satellites and ground-based instruments, *Journal of Geophysical Research: Atmospheres*, 122,



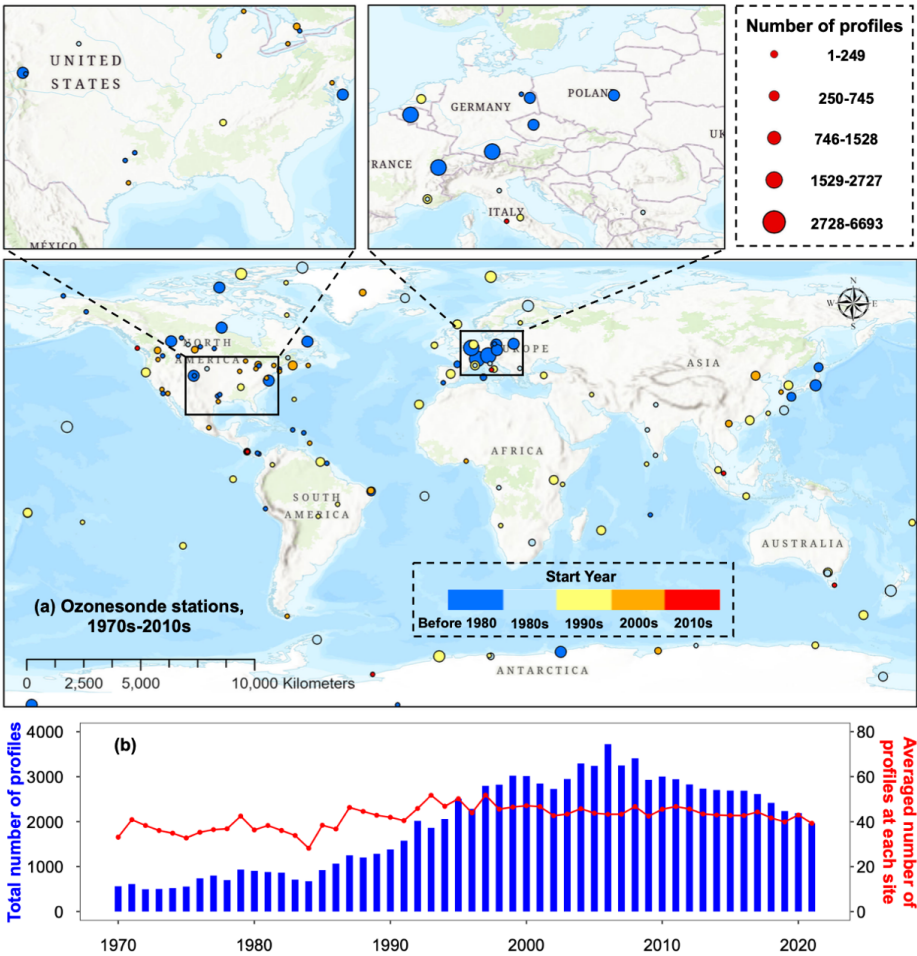
- 872 13,000-013,025, 2017.
- 873 Thouret, V., Marenco, A., Logan, J. A., Nédélec, P., and Grouhel, C.: Comparisons of ozone
874 measurements from the MOZAIC airborne program and the ozone sounding network at eight
875 locations, *Journal of Geophysical Research: Atmospheres*, 103, 25695-25720, 1998.
- 876 Tilmes, S., Lamarque, J.-F., Emmons, L., Conley, A., Schultz, M., Saunio, M., Thouret, V.,
877 Thompson, A., Oltmans, S., and Johnson, B.: Ozonesonde climatology between 1995 and
878 2011: description, evaluation and applications, *Atmospheric Chemistry and Physics*, 12,
879 7475-7497, 2012.
- 880 Van Malderen, R., Allaart, M. A., De Backer, H., Smit, H. G., and De Muer, D.: On instrumental
881 errors and related correction strategies of ozonesondes: possible effect on calculated ozone
882 trends for the nearby sites Uccle and De Bilt, *Atmospheric Measurement Techniques*, 9, 3793-
883 3816, 2016.
- 884 Wang, G., Kong, Q., Xuan, Y., Wan, X., Chen, H., and Ma, S.: Development and application of
885 ozonesonde system in China, *Advances in Earth Science*, 18, 471, 2003.
- 886 Wang, H. J., Cunnold, D. M., Thomason, L. W., Zawodny, J. M., and Bodeker, G. E.: Assessment
887 of SAGE version 6.1 ozone data quality, *Journal of Geophysical Research: Atmospheres*, 107,
888 ACH 8-1-ACH 8-18, 2002.
- 889 Wild, O., Voulgarakis, A., O'Connor, F., Lamarque, J.-F., Ryan, E. M., and Lee, L.: Global
890 sensitivity analysis of chemistry–climate model budgets of tropospheric ozone and OH:
891 exploring model diversity, *Atmospheric Chemistry and Physics*, 20, 4047-4058, 2020.
- 892 Witte, J. C., Thompson, A. M., Schmidlin, F., Northam, E. T., Wolff, K. R., and Brothers, G. B.:
893 The NASA Wallops Flight Facility digital ozonesonde record: Reprocessing, uncertainties,
894 and dual launches, *Journal of Geophysical Research: Atmospheres*, 124, 3565-3582, 2019.
- 895 Witte, J. C., Thompson, A. M., Smit, H. G., Vömel, H., Posny, F., and Stübi, R.: First reprocessing
896 of Southern Hemisphere ADDitional OZonesondes profile records: 3. Uncertainty in ozone
897 profile and total column, *Journal of Geophysical Research: Atmospheres*, 123, 3243-3268,
898 2018.
- 899 Witte, J. C., Thompson, A. M., Smit, H. G., Fujiwara, M., Posny, F., Coetzee, G. J., Northam, E.
900 T., Johnson, B. J., Sterling, C. W., and Mohamad, M.: First reprocessing of Southern
901 Hemisphere ADDitional OZonesondes (SHADOZ) profile records (1998–2015): 1.
902 Methodology and evaluation, *Journal of Geophysical Research: Atmospheres*, 122, 6611-



- 903 6636, 2017.
- 904 Worden, J., Liu, X., Bowman, K., Chance, K., Beer, R., Eldering, A., Gunson, M., and Worden,
905 H.: Improved tropospheric ozone profile retrievals using OMI and TES radiances,
906 Geophysical Research Letters, 34, 2007.
- 907 Xu, W., Xu, X., Lin, M., Lin, W., Tarasick, D., Tang, J., Ma, J., and Zheng, X.: Long-term trends
908 of surface ozone and its influencing factors at the Mt Waliguan GAW station, China–Part 2:
909 The roles of anthropogenic emissions and climate variability, Atmospheric Chemistry and
910 Physics, 18, 773–798, 2018.
- 911 Xuan, Y., Ma, S., Chen, H., Wang, G., Kong, Q., Zhao, Q., and Wan, X.: Intercomparisons of
912 GPSO3 and Vaisala ECC ozone sondes, Plateau Meteorology, 23, 394–399, 2004.
- 913 Young, P. J., Naik, V., Fiore, A. M., Gaudel, A., Guo, J., Lin, M. Y., Neu, J. L., Parrish, D. D.,
914 Rieder, H. E., Schnell, J. L., Tilmes, S., Wild, O., Zhang, L., Ziemke, J., Brandt, J., Delcloo,
915 A., Doherty, R. M., Geels, C., Hegglin, M. I., Hu, L., Im, U., Kumar, R., Luhar, A., Murray,
916 L., Plummer, D., Rodriguez, J., Saiz-Lopez, A., Schultz, M. G., Woodhouse, M. T., and Zeng,
917 G.: Tropospheric Ozone Assessment Report: Assessment of global-scale model performance
918 for global and regional ozone distributions, variability, and trends, Elementa: Science of the
919 Anthropocene, 6, 10.1525/elementa.265, 2018.
- 920 Zbinden, R., Thouret, V., Ricaud, P., Carminati, F., Cammas, J.-P., and Nédélec, P.: Climatology
921 of pure tropospheric profiles and column contents of ozone and carbon monoxide using
922 MOZAIC in the mid-northern latitudes (24° N to 50° N) from 1994 to 2009, Atmospheric
923 Chemistry and Physics, 13, 12363–12388, 2013.
- 924 Zeng, G., Morgenstern, O., Williams, J. H., O'Connor, F. M., Griffiths, P. T., Keeble, J., Deushi,
925 M., Horowitz, L. W., Naik, V., and Emmons, L. K.: Attribution of stratospheric and
926 tropospheric ozone changes between 1850 and 2014 in CMIP6 models, Journal of
927 Geophysical Research: Atmospheres, 127, e2022JD036452, 2022.
- 928 Zeng, Y., Zhang, J., Li, D., Liao, Z., Bian, J., Bai, Z., Shi, H., Xuan, Y., Yao, Z., and Chen, H.:
929 Vertical distribution of tropospheric ozone and its sources of precursors over Beijing: Results
930 from~ 20 years of ozonesonde measurements based on clustering analysis, Atmospheric
931 Research, 284, 106610, 2023.
- 932 Zhang, J., Li, D., Bian, J., Xuan, Y., Chen, H., Bai, Z., Wan, X., Zheng, X., Xia, X., and Lü, D.:
933 Long-term ozone variability in the vertical structure and integrated column over the North



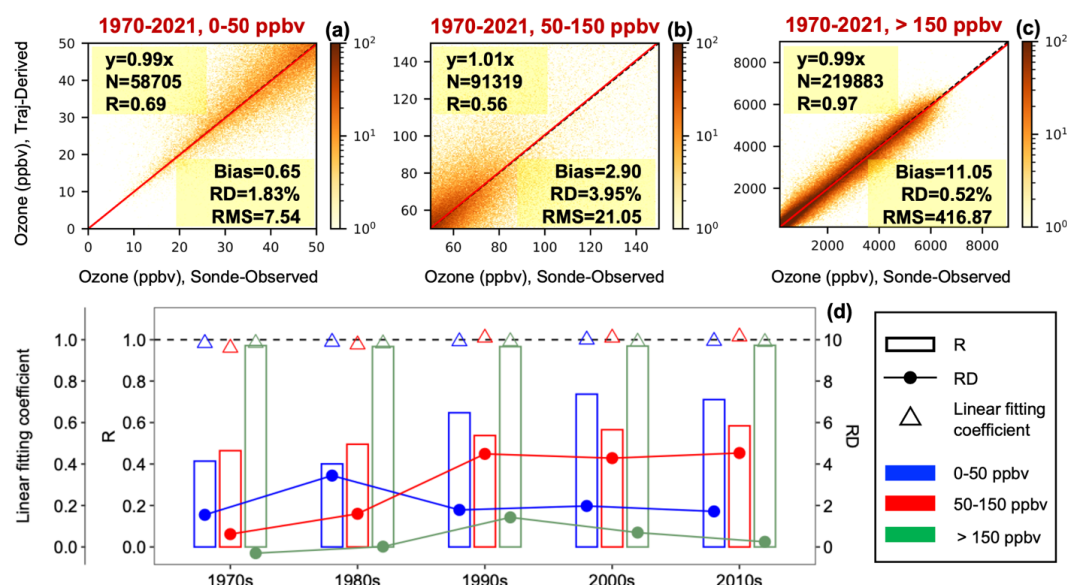
934 China Plain: Results based on ozonesonde and Dobson measurements during 2001–2019,
935 Environmental Research Letters, 16, 074053, 2021.
936



937
938
939 Figure 1. (a) Global distribution of ozonesonde stations used in this study to construct TOST-v2.
940 Station details are provided in Table S1. The size and color of the dots indicate the total number
941 of sounding profiles and the start year of the measurement time series. (b) The total number of
942 profiles per year (left y-axis, blue bars) and the average number of profiles per site and per year

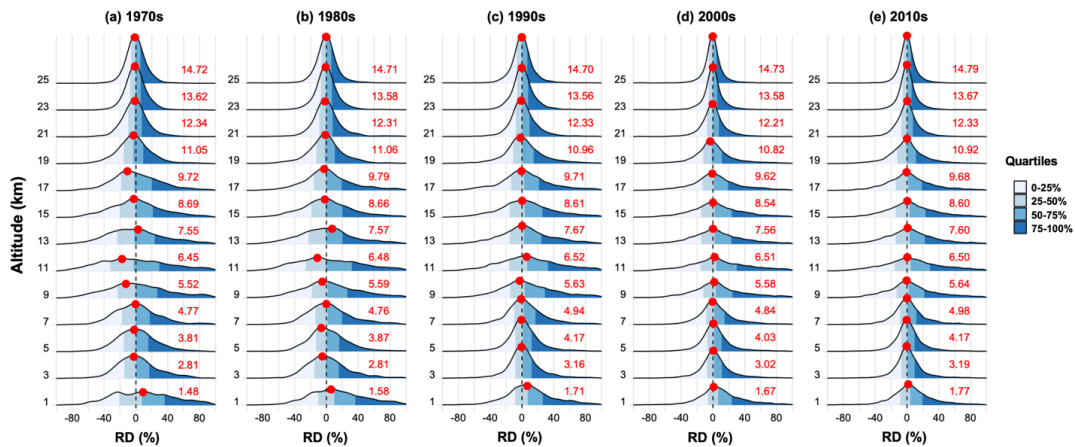


(right y-axis, red dots and line) from 1970 to 2021.



944

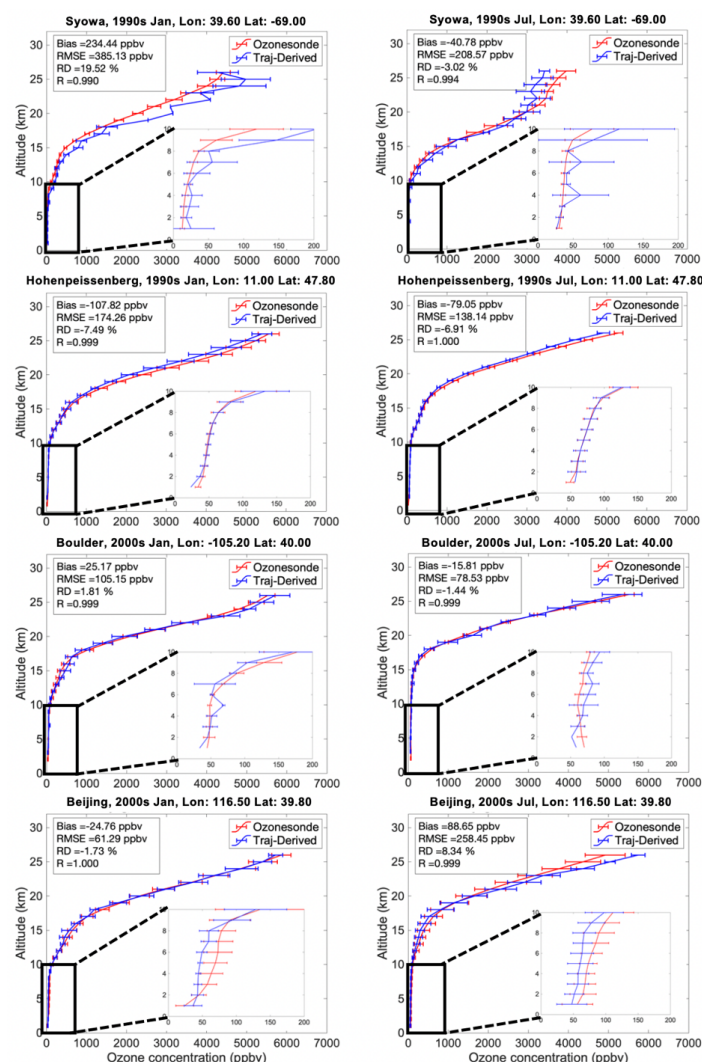
Figure 2. (a-c) Comparison of monthly average tropospheric ozone mixing ratios from ozonesondes (Sonde-Observed) and trajectory-derived TOST data (Traj-Derived) for the entire study period of ozone concentration at 0-50 ppbv, 50-150 ppbv and >150 ppbv. Solid red lines represent the linear fitting line (with the intercept set to 0) and dashed black lines denote the 1:1 axis. N is the total number of data points, R is the correlation coefficient, Bias is the overall average difference in monthly mean values [Traj-Derived ozone - Sonde-Observed ozone, in ppbv], RD is the relative difference in % [$100 \times (\text{Traj-Derived ozone} - \text{Sonde-Observed ozone}) / \text{Sonde-Observed ozone}$], and RMS is the root mean square difference in ppbv). Note that Traj-Derived ozone at each station is derived without input from the station itself; that is, Traj-Derived represents an ensemble of 141 separate computations of TOST, each one withholding a single validation station. (d) the R (bars), RD (dots and lines) and linear fitting coefficient (with the intercept set to 0; triangles) between the Traj-Derived ozone and Sonde-Observed ozone by decade. The dashed line denotes where the linear fitting coefficient is 1.



958

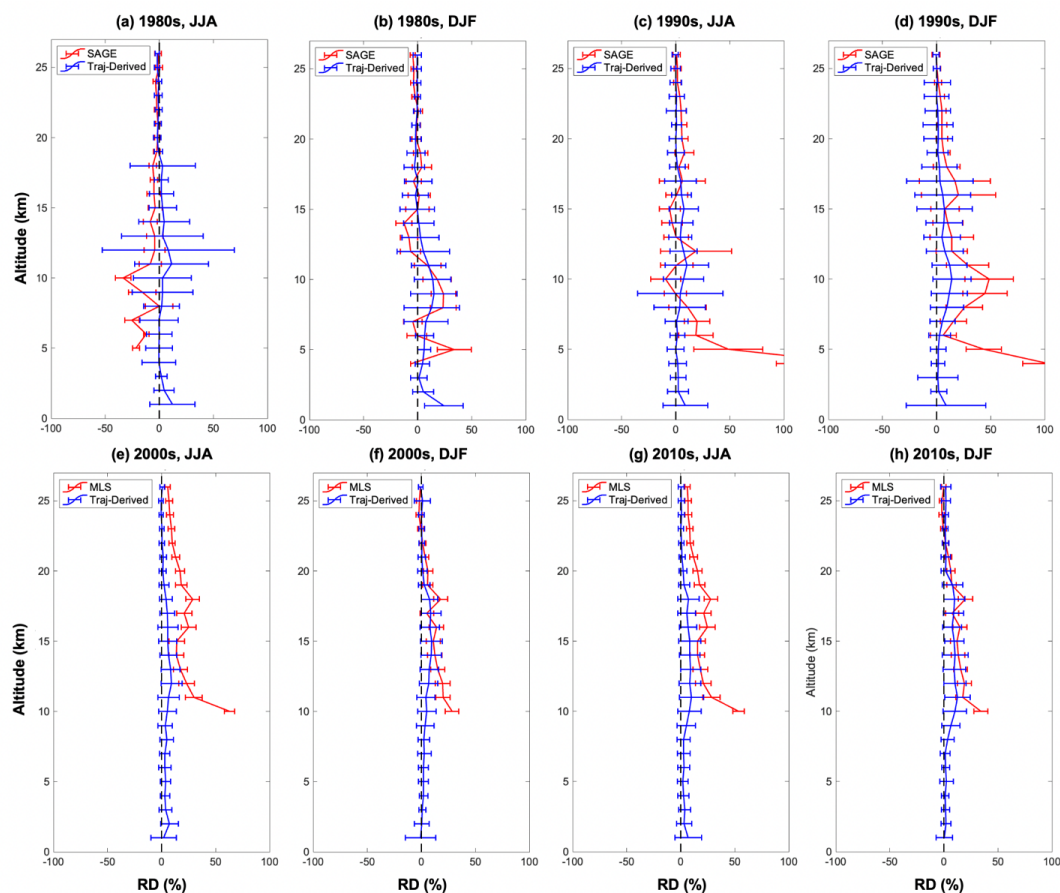
959

960 Figure 3. The relative difference (RD) of the monthly ozone mixing ratios between ozonesonde
961 and Traj-Derived data by altitude in the 1970s, 1980s, 1990s, 2000s and 2010s, respectively. The
962 frequency distribution of RD at every other altitudes is shown (y-axis: frequency in %, x-axis: RD
963 in %), with the colors denoting the 4 quartiles of RD. The dashed line indicates zero difference in
964 RD. The red dot and number represent the maximum frequency and the corresponding frequency
965 value in %, respectively.



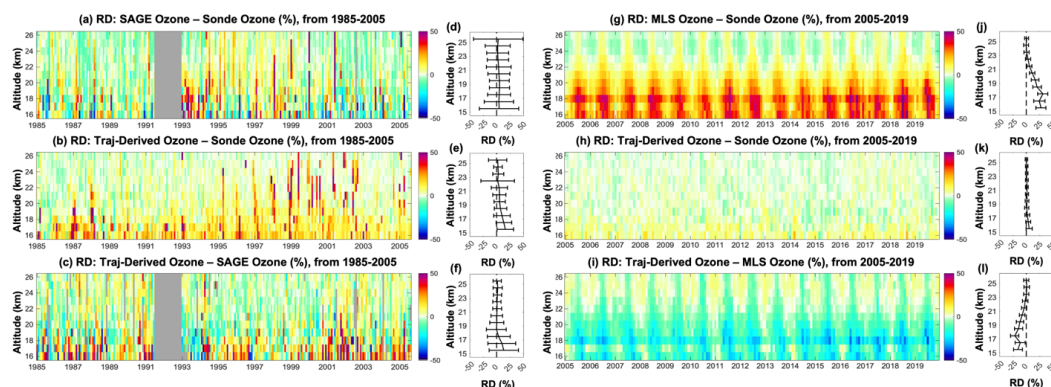
966

967 Figure 4. Decadal monthly mean ozone profiles at Syowa and Hohenpeissenberg in January and
 968 July 1990s, and at Boulder and Beijing in January and July 2000s. The red line denotes ozonesonde
 969 ozone and the blue line denotes trajectory-derived ozone without the input from the station itself.
 970 The error bar is ± 2 times the standard error of the mean (equivalent to 95 % confidence limits on
 971 the averages). To better compare the difference of ozone profiles in the troposphere, a zoom-in
 972 window from 0-10 km is provided in each sub-figure.



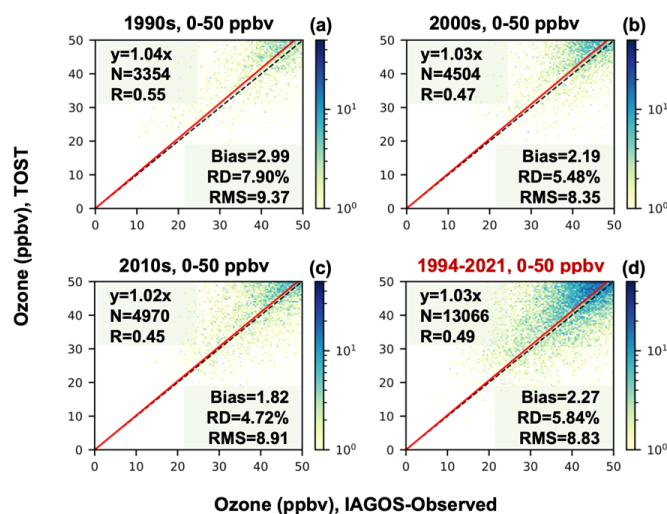
973

974 Figure 5. (a-d) Mean relative difference (RD) of the monthly ozone mixing ratios between the Traj-
 975 Derived and ozonesonde data (blue line) and between the SAGE and ozonesonde data (red line) in
 976 JJA (June-July-August) and DJF (December-January-February), in the 1980s and 1990s. (e-h)
 977 Decadal seasonal mean RD between the trajectory-derived and ozonesonde data (blue line) and
 978 between the MLS and ozonesonde data (red line) in JJA and DJF, in the 2000s and 2010s. The
 979 error bars represent ± 1 standard deviation of the seasonal mean RDs at each altitude in each decade.



980

981 Figure 6. (a). The relative difference (RD) between ozonesonde and SAGE ozone data in each
 982 month and at each altitude during 1985-2005 over 16-26 km [$RD = 100 \times (\text{SAGE ozone} -$
 983 $\text{ozonesonde ozone}) / \text{ozonesonde ozone}$, in %]. The mean RD over 1985-2005 at each level is shown
 984 on the right, where the error bars represent the standard deviation of the monthly RD over 1985-
 985 2005. Note that the Pinatubo-affected SAGE profiles are excluded during July 1991- December
 986 1992 (filled with gray color). (b) same as (a), but for the RD between ozonesondes and Traj-derived
 987 data, [$RD = 100 \times (\text{Traj-derived ozone} - \text{ozonesonde ozone}) / \text{ozonesonde ozone}$, in %]. (c) same as
 988 (a), but for the RD between Traj-derived and SAGE ozone data [$RD = 100 \times (\text{Traj-derived ozone}$
 989 $- \text{SAGE ozone}) / (0.5 \times \text{Traj-derived ozone} + 0.5 \times \text{SAGE ozone})$, in %]. (d-f) the averaged RD by
 990 altitude corresponding to (a-c). (g-l) same as (a-f), but for the period of 2005-2019 and the satellite
 991 measurements are from MLS ozone.



992

993 Figure 7. The comparison of monthly ozone mixing ratios between IAGOS-observed (x-axis
994 labeled: IAGOS-Observed) and TOST data (y-axis labelled: TOST) by decade (a-c) and for the
995 entire study period (d) of ozone concentration at 0-50 ppbv. Solid red lines represent the linear
996 fitting line (with the intercept set to 0) and dashed black lines denote the 1:1 axis. N is the total
997 number of data points, R is the correlation coefficient (unitless), Bias is the difference in monthly
998 mean values [TOST ozone - IAGOS ozone, unit: ppbv], RD is the relative difference [$100 \times (\text{TOST}$
999 ozone - IAGOS ozone)/(0.5 $\times$ TOST ozone + 0.5 $\times$ IAGOS ozone)], and RMS the root mean square
1000 difference (unit: ppbv).

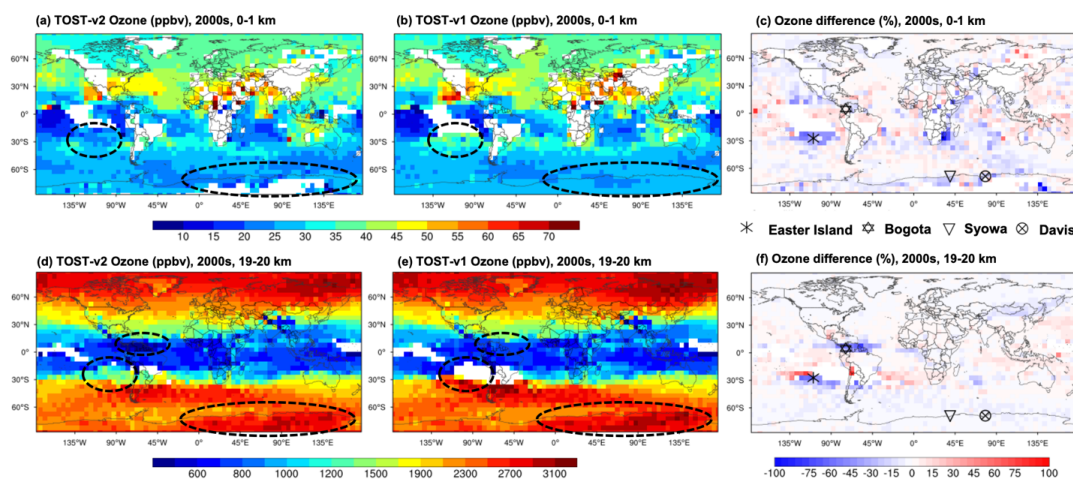
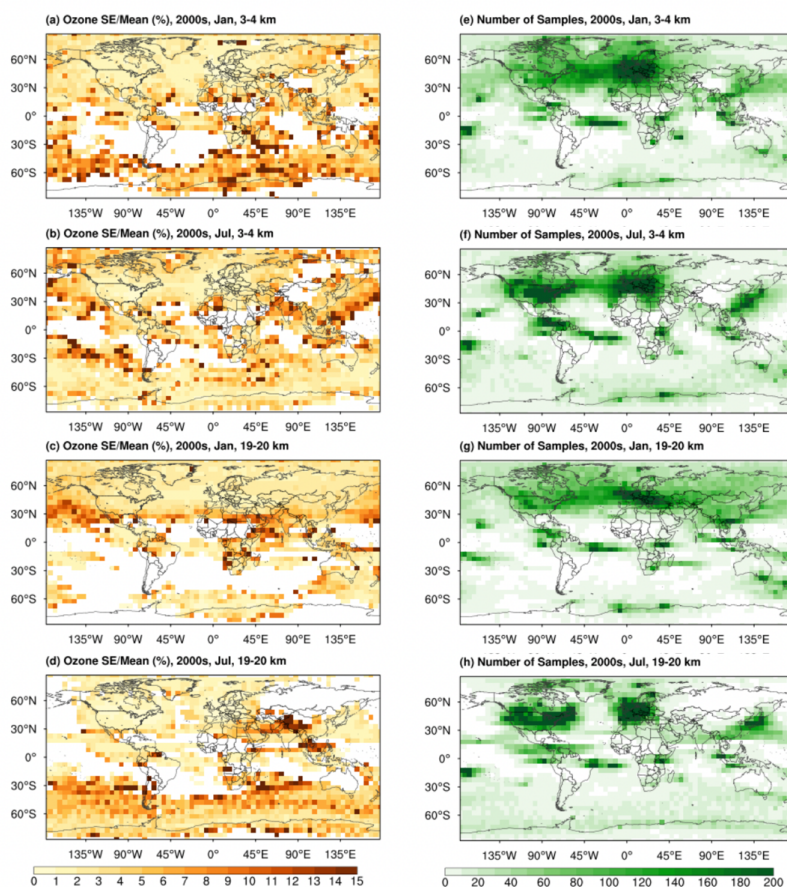


Figure 8. (a, b) The global distributions of ozone in TOST-v2 (a) and TOST-v1 (b) over 0-1 km in the 2000s. (d, e) The same as for (a, b), but over 19-20 km. The dashed circles indicate regions with large differences between the two versions. (c) The global distributions of RD between TOST-v2 and TOST-v1 [$RD = 100 \times (TOST-v2 - TOST-v1) / (0.5 \times TOST-v2 + 0.5 \times TOST-v1)$, in %] over 0-1 km in the 2000s. (f) The same as for (c), but over 19-20 km. The markers indicate the positions of Davis, Easter Island, Bogota and Syowa stations.



1008

1009

1010 Figure 9. (a-d) Global distribution of the standard error of the mean (left panels, in %) for the
1011 decadal monthly mean ozone in January and July 2000s at 3-4 km (a and b) and 19-20 km (c and
1012 d). (e-h) the same as (a-d), but for the number of samples in each $5 \times 5^\circ$ bin.

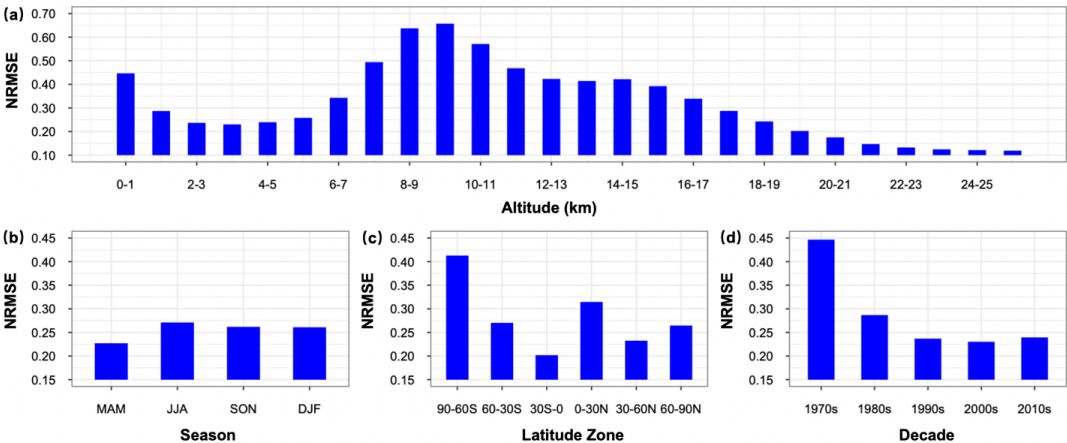
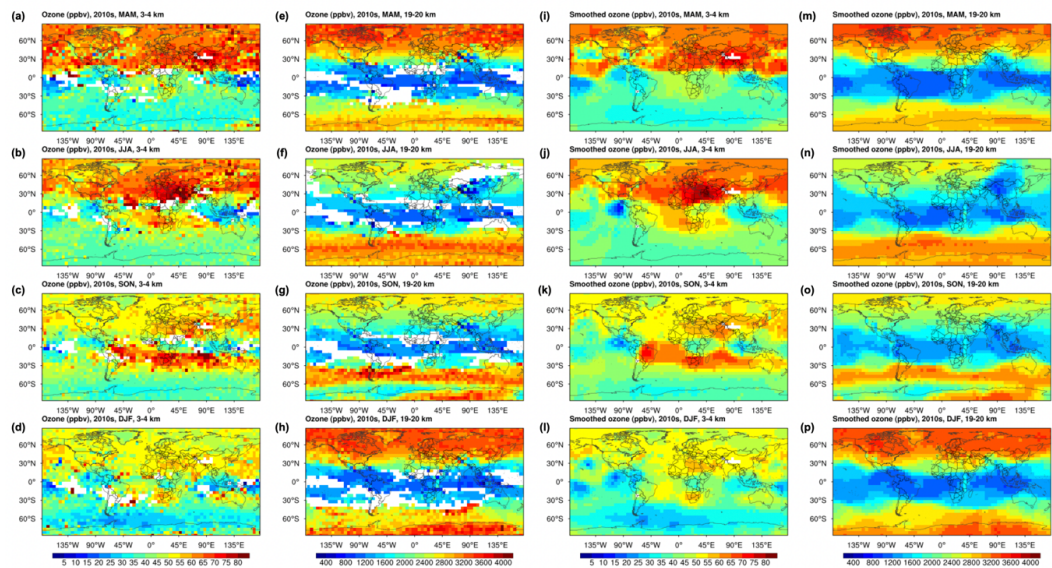


Figure 10. The Normalized Root Mean Squared Error (NRMSE, unitless) of TOST over 1990-2021 by altitude, and the average NRMSE over all altitudes by (a), season (b), latitudinal zone (c), and decade (d). The NRMSE is calculated as the RMS difference of monthly ozone mixing ratio between ozonesondes and Traj-Derived ozone divided by the mean ozone mixing ratio from ozonesondes measurements.



1020

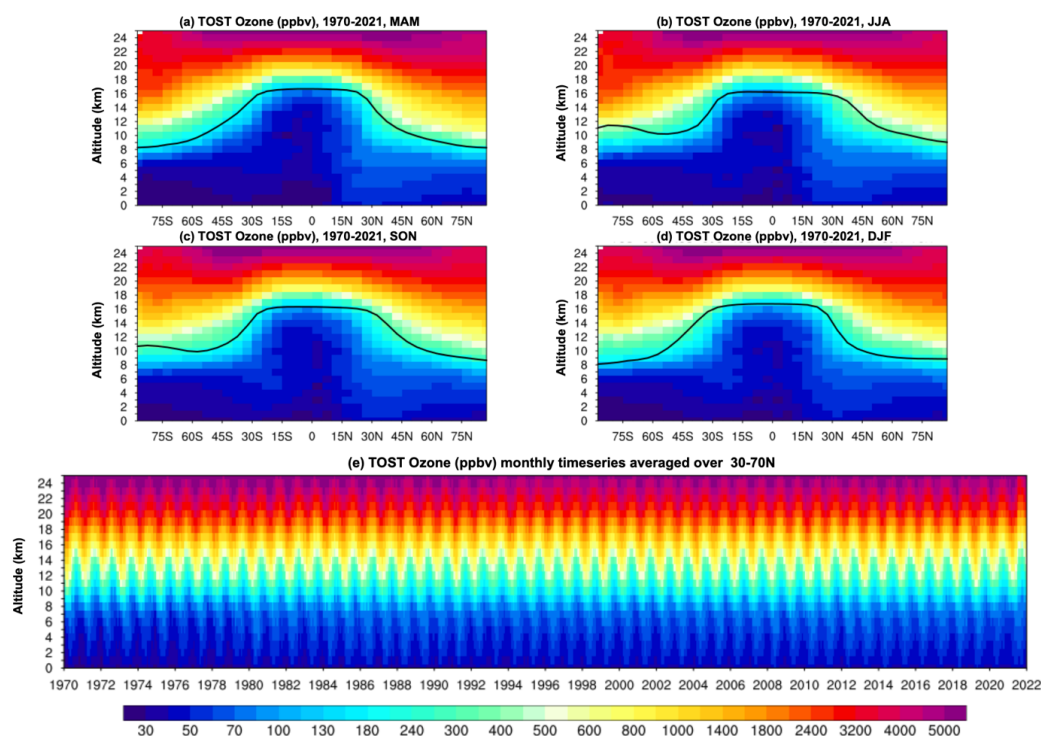


1021

1022 Figure 11. Global distribution of decadal mean TOST ozone at 3-4 km and 19-20 km in MAM
1023 (March-April-May), JJA (June-July-August), SON (September-October-November) and DJF
1024 (December-January-February) in the 2010s (a-h), and the corresponding smoothed TOST ozone
1025 (i-p).



1026



1027

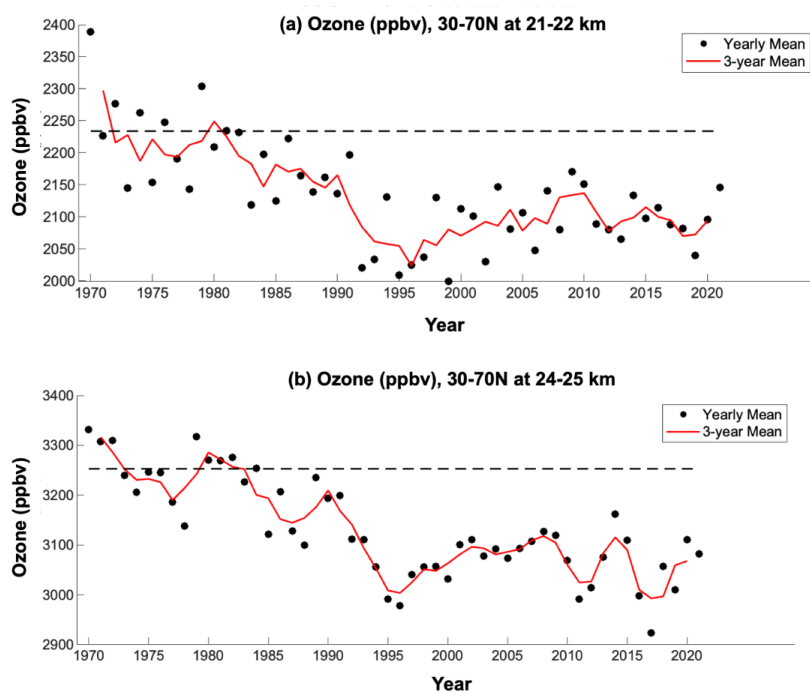
1028

1029 Figure 12. (a-d) The latitude-altitude distribution of TOST ozone averaged over 1970-2021 in each

1030 season. The solid black lines represent the mean tropopause height over 1970-2021 in each season.

1031 (e) time series of the monthly mean TOST ozone over 30-70°N at each altitude level from 1970 to

1032 2021.



1033

1034 Figure 13. TOST time series of the annual mean ozone mixing ratios averaged over 30-70°N over
 1035 21-22 km altitude (a) and 24-25 km altitude (b). The black dots represent the annual mean ozone
 1036 concentrations from the area-weighted average of the gridpoints over 30-70°N with ozone data
 1037 throughout 1970-2021. The red line is the 3-year running mean. The black dashed line indicates
 1038 the average ozone concentrations in the 1970s.