Gravitational separation of Ar/N₂ and age of air in the lowermost stratosphere in airborne observations and a chemical transport model

Benjamin Birner[1], Martyn P. Chipperfield[2,3], Eric J. Morgan[1], Britton B. Stephens[4], Marianna Linz[5], Wuhu Feng[2,6], Chris Wilson[2,3], Jonathan D. Bent[1,7], Steven C. Wofsy[5], Jeffrey Severinghaus[1], Ralph F. Keeling[1]

[1]Scripps Institution of Oceanography, UC San Diego, La Jolla, CA 92093, USA
[2]School of Earth and Environment, University of Leeds, Leeds, LS2 9JT, UK
[3]National Centre for Earth Observation, University of Leeds, Leeds, LS2 9JT, UK
[4]National Center for Atmospheric Research, Boulder, CO 80301, USA
[5]Department of Earth and Planetary Sciences, and School of Engineering and Applied Sciences, Harvard University, Cambridge, MA 02138, USA
[6]National Centre for Atmospheric Science, University of Leeds, Leeds, LS2 9JT, UK
[7]now at: Picarro, Inc., Santa Clara, CA 95054, USA

Correspondence to: Benjamin Birner (bbirner@ucsd.edu)

Abstract. Accurate simulation of atmospheric circulation, particularly in the lower stratosphere, is challenging due to unresolved wave-mean flow interactions and limited high-resolution observations for validation. Gravity-induced pressure gradients lead to a small but measurable separation of heavy and light gases by molecular diffusion in the stratosphere. Because the relative abundance of Ar to N₂ is exclusively controlled by physical transport, the argon-to-nitrogen ratio (Ar/N₂) provides an additional constraint on circulation and the age of air (AoA), i.e. the time elapsed since entry of an air parcel into the stratosphere. Here we use airborne measurements of N₂O and Ar/N₂ from nine campaigns with global coverage spanning 2008–2018 to calculate AoA and to quantify gravitational separation in the lowermost stratosphere. To this end, we develop a new N₂O-AoA relationship using a Markov Chain Monte Carlo algorithm. We observe that gravitational separation increases systematically with increasing AoA for samples with AoA between 0 to 3 years. These observations are compared to a simulation of the TOMCAT/SLIMCAT 3-D chemical transport model, which has been updated to include gravitational fractionation of gases. We demonstrate that although AoA at old ages is slightly underestimated in the model, the relationship between Ar/N₂ and AoA is robust and agrees with the observations. This highlights the potential of Ar/N₂ to become a new AoA tracer that is subject only to physical transport phenomena and can supplement the suite of available AoA indicators.

1. Introduction

30 Transport in the middle atmosphere is driven by a combination of advection by the Brewer-Dobson circulation (BDC) (Brewer, 1949; Dobson, 1956) and quasi-horizontal, two-way mixing by breaking waves (Holton et al., 1995). Models consistently

predict an acceleration of the BDC due to climate change (Butchart, 2014) but subgrid-scale mixing processes and momentum transfer by unresolved buoyancy waves limit our ability to accurately simulate circulation in the stratosphere (Haynes, 2005; Plumb, 2007). An acceleration of the BDC has important repercussions for stratosphere-troposphere exchange (STE), and thus recovery of the ozone layer and the greenhouse effect of stratospheric water vapor, however observational evidence for an acceleration of the BDC is weak (Engel et al., 2009, 2017; Waugh, 2009; Ray et al., 2010, 2014).

The mean age of air (AoA) is a widely used indicator of stratospheric circulation (Hall and Plumb, 1994; Waugh and Hall, 2002; Linz et al., 2016). Air can be transported to any location r in the stratosphere via a myriad of different paths, and each path will have an associated transit time. The probability density function that describes the likelihood for air to reach location r with a specific transit time is called the age spectrum. Although the age spectrum is not directly observable, some aspects of its shape can be inferred from observations of short- and long-lived tracers (Andrews et al., 1999, 2001; Schoeberl et al., 2005; Hauck et al., 2019, 2020; Podglajen and Ploeger, 2019). For tracers that are conserved in the stratosphere and whose concentrations increase approximately linearly with time in the troposphere, such as SF_6 and CO_2, the mean AoA, i.e., the first moment of the distribution, can simply be calculated as the time difference, or "lag time", to when tracer concentrations in the upper troposphere last had comparable values as measured in a stratospheric sample (Hall and Plumb, 1994; Boering et al., 1996; Waugh and Hall, 2002). The stratospheric concentration of N_2O has also been calibrated as an independent AoA tracer by relating the gradual photolytic loss of N_2O in the stratosphere to AoA as determined from CO_2 (Boering et al., 1996; Andrews et al., 2001; Linz et al., 2017).

In contrast to early measurements made on rocket samples (Bieri and Koide, 1970), Ishidoya et al. (2008, 2013) have shown using a balloon-borne sampling system that stratospheric air is detectably fractionated by gravitational settling (GS), with the degree of fractionation strongly correlated to AoA (Ishidoya et al., 2008, 2013, 2018; Sugawara et al., 2018; Belikov et al., 2019). GS leads to depletion of heavier gases in the stratosphere yielding lower ratios of heavy to light gases, e.g. Ar/N_2, $^{18}O/^{16}O$ of O_2 and $^{15}N/^{14}N$ of N_2, with increasing elevation. The vertical gradients induced by gravimetric separation are steeper at higher latitudes (Ishidoya et al., 2008, 2013, 2018; Sugawara et al., 2018), and consistent with patterns observed in stratospheric AoA (Sugawara et al., 2018; Belikov et al., 2019). Gravimetric settling in the stratosphere has also been simulated in 1-D (Keeling, 1988; Ishidoya et al., 2008), 2-D (Ishidoya et al., 2013, 2018; Sugawara et al., 2018), and 3-D stratospheric models (Belikov et al., 2019). 2-D and 3-D models both show a pattern of GS which increases with altitude and latitude, similar to the patterns observed in tracers with a significant stratospheric sink such as N_2O, and consistent with a positive correlation with AoA (Ishidoya et al., 2013; Sugawara et al., 2018; Belikov et al., 2019). Here we attempt to calibrate GS of Ar/N_2 as an AoA tracer similar to previous work on the N_2O-AoA relationship.

Observing GS in the stratosphere is challenging, however, as the signals are small and because of the need to avoid artifacts caused by temperature- and pressure-induced fractionation near the sampling inlet (Blaine et al., 2006; Ishidoya et al., 2008,

2013). The resulting scatter in existing balloon-based measurements precludes a clear evaluation of the relationship between GS and AoA (Belikov et al., 2019).

65 Here we present a dataset of gravitational fractionation of Ar/N_2 and AoA observations made on flask samples from three airborne research projects, from 9 campaigns in the lowermost stratosphere with mean AoA <3 years (Wofsy et al., 2012, 2018; Stephens, 2017). The HIAPER Pole-to-Pole Observations (HIPPO) project was a global survey of the Pacific troposphere to lower stratosphere on the NSF/NCAR Gulfstream-V aircraft (Wofsy et al., 2011), composed of 5 individual campaigns from 2008–2011. The O_2/N_2 Ratio and CO_2 Airborne Southern Ocean (ORCAS) study was conducted using the

70 same aircraft but focused on the Drake Passage and Antarctic Peninsula during Jan–Feb 2016 (Stephens et al., 2018). The Atmospheric Tomography (ATom) project was a survey of the troposphere and lower stratosphere of the Pacific and Atlantic Ocean basins on the NASA DC-8 aircraft, composed of 4 individual campaigns from 2016–2018 (Prather et al., 2017). The observations are compared to new simulations of GS with the TOMCAT/SLIMCAT (Chipperfield, 2006) tracer transport model. Our goals are to demonstrate the consistency of our data with gravitational fractionation, to evaluate model

75 performance, and to highlight the potential of Ar/N_2 as a new age tracer.

2. Methods

2.1. Measurements

Discrete 1.5L flask samples were taken with the NCAR/Scripps Medusa Whole Air Sampler (Bent, 2014) (https://www.eol.ucar.edu/instruments/ncar-scripps-airborne-flask-sampler). Medusa holds 32 borosilicate glass flasks sealed

80 with Viton o-rings and uses active pressure control to fill the flasks with cryogenically dried air to ~760 torr. Flasks are shipped to Scripps Institution of Oceanography (SIO) for analysis of Ar/N_2 ratios on an IsoPrime Mass Spectrometer. We report changes in Ar/N_2 ratios in delta notation

$$\delta = \left(\frac{R_{SA}}{R_{REF}} - 1\right) \times 10^6 (per\ meg) \qquad 1$$

where R_{SA} is the mixing or isotope ratio in the sample and R_{REF} the ratio in a reference mixture. Measurements are made on the Scripps O2 Program Argon Scale, as defined on 21 Jan 2020. δ(Ar/N_2) values are reported after applying an offset to the

85 data to yield a mean of zero in the tropical airborne observations of the free troposphere between 3-8 km.

Uncertainty in δ(Ar/N_2) observations arises from a combination of analytical limitations and artifactual fractionation during sampling. Replicate agreement of surface flasks shows a 1σ repeatability of ±6.1 per meg for δ(Ar/N_2) but additional scatter in the data may be introduced by small leaks in the Medusa system and thermal or pressure gradient fractionation at the sample intake (Morgan et al., 2020). For further details, see Bent (2014). These effects are challenging to separate from true

90 atmospheric variability and differ between campaigns because sampling strategies have generally improved over time. To

constrain uncertainty due to aircraft sampling, we quantify total variability in observations from the presumably fairly homogeneous tropical troposphere between 3-8 km (Fig. S1). Data from the earlier campaigns HIPPO 1–5 show a pooled standard deviation of 24 per meg whereas data from ATom 2–4 yield a pooled standard deviation of 9 per meg, illustrating advances in sampling and sample handling. While no ORCAS samples are available in the tropical troposphere, ORCAS data show similar scatter as ATom 2–4 data between 20°S and 50°S. For all campaigns, the uncertainty is small compared to the stratospheric signal of tens to hundreds per meg shown below. We also show that the stratospheric signal in δ(Ar/N$_2$) is not due to pressure-dependent inlet fractionation by evaluating the residual δ(Ar/N$_2$), which has been corrected for the natural gravitational signal (Fig. S2).

We have compiled available simultaneous, high frequency measurements of a range of other trace gases including N$_2$O, CO$_2$, O$_3$, CH$_4$, and CO to identify Medusa samples with stratospheric influence and calculate mean AoA. N$_2$O was measured continuously with a precision of 0.09 ppb at 1 Hz frequency using the Harvard Quantum Cascade Laser Spectrometer (QCLS) (Santoni et al., 2014) during HIPPO and ORCAS, and measured every 1–3 min using the NOAA gas chromatograph PAN and Trace Hydrohalocarbon ExpeRiment (PANTHER) during ATom. The Unmanned Aircraft Systems Chromatograph for Atmospheric Trace Species (UCATS) was used to measure O$_3$ during HIPPO and ATom. O$_3$ was not measured on ORCAS. We use continuous H$_2$O data from the NCAR open path near-infrared multi-pass spectrometer Vertical Cavity Surface Emitting Laser (VCSEL) (Zondlo et al., 2010) for HIPPO and ORCAS whereas H$_2$O was measured using the NASA Diode Laser Hygrometer (DLH) (Diskin et al., 2002) during ATom flights. CH$_4$, CO$_2$, and CO were measured by QCLS during HIPPO and ORCAS, and by the NOAA Picarro (Karion et al., 2013) during ATom. An averaging kernel is applied to the continuous and semi-continuous aircraft data, such as N$_2$O, O$_3$, and H$_2$O, to match it to Medusa samples. The kernel multiplies a weighting function $w_i(t)$ with all continuous data before time t_s when Medusa switched from sample flask i to sample flask $i+1$. $w_i(t)$ for each sample i is given by

$$w_i(t) = \exp\left[\frac{-t_s - t}{\tau}\right], \hspace{4cm} 2$$

where t is each 1s-increment of the continuous data and $\tau = \frac{V}{Q}$ is the flushing time of air in a Medusa flask determined by the flask volume V and airflow Q.

Stratospheric samples are identified based on their N$_2$O, O$_3$, and water vapor levels. Classification based on chemical composition rather than potential temperature or altitude effectively selects samples with a clear stratospheric signature in the lowermost stratosphere and excludes air which has experienced substantial mixing with tropospheric air masses. We label samples as "stratospheric" if (i) water vapor levels are below 15 ppm and either (ii.a) O$_3$ values exceed 140 ppb or (ii.b) N$_2$O (detrended to a reference year of 2009) is below 315 ppb. These criteria yield 234 lower stratospheric samples with high quality N$_2$O and δ(Ar/N$_2$) data, spanning a wide range of latitudes poleward of 40° in both the Northern and Southern Hemisphere (Fig. 1). We use Medusa samples from all 5 HIPPO campaigns, ORCAS, and ATom 2–4. We do not use samples from ATom

1 because inlet fractionation due to the unique inlet design and location on the DC-8 on this campaign introduced apparent biases on the order of 30 per meg. An additional 25 stratospheric samples are available from the START-08 campaign on the NSF/NCAR GV, but we have not used these here because of the limited number of flasks collected and greater scatter in the data.

2.2. AoA calculation

Stratospheric (mean) AoA for Medusa samples is calculated from N_2O using an updated hemisphere-specific N_2O-AoA relationship. Our method broadly follows Andrews et al. (2001) who assumed a bimodal age spectrum and used multiple observations of CO_2 binned by N_2O values to resolve the seasonal cycle of CO_2 in each bin. Properties of the age spectrum for each N_2O bin, including mean AoA were constrained by optimizing the agreement between observed CO_2 concentrations and concentration implied by randomly generated age spectra in each N_2O bin. Andrews et al. (2001) used a highly efficient "genetic algorithm" to yield the most likely relationship between mean AoA and N_2O in each bin. In contrast, we use a more computationally costly Markov chain Monte Carlo (MCMC) method that allows us to obtain more robust uncertainties for all estimated parameters of the age spectrum. Following Malinverno (2002) and Green (1995), our algorithm builds on a Metropolis-Hasting sampler (Metropolis et al., 1953; Hastings, 1970) to evaluate probability distributions for each age spectrum parameter and automatically choses whether a unimodal or bimodal representation of the age spectrum is more appropriate in each N_2O bin. Finally, we constructed a new tropical upper troposphere reference time series for CO_2 in this study to ensure maximum consistency between all observations used. Analytical and methodological uncertainties are propagated thoroughly and reported as the 95% confidence interval around a mean.

We consider the tropical upper troposphere as the single entry-point of air into the stratosphere for our AoA-calculation. Mixing between tropospheric and lower stratospheric air can occur outside the tropics leading to air masses that are transitional between stratospheric and tropospheric in the lowermost extratropical stratosphere (e.g., Škerlak et al., 2014). This additional pathway for tropospheric entry has been accounted for in previous studies via additional model information and/or multiple tracers (Ray et al., 1999; Hoor et al., 2004; Bönisch et al., 2009; Hauck et al., 2020). However, we are limited by the number of tracers available to us, and therefore instead aim to largely exclude samples influenced through this additional pathway by using threshold in N_2O, O_3, and H_2O when identifying stratospheric samples. Because of this sample selection strategy, we note that the ages calculated here should not be considered representative for the region that air was sampled in.

Our upper troposphere reference time series (TRTS) consists of a long-term trend and a representation of the mean seasonal cycle. Because direct CO_2 observations in the region are limited, the long-term trend and a first guess of the mean seasonal cycle are estimated from monthly mean surface observations at the Mauna Loa station in Hawaii (MLO, 19°N, 155°W, Fig 2, panel (a)) (Keeling et al., 2001) which are later adjusted to match airborne observations in the tropics (20°S-20°N) above 8 km (Fig. 2, panel (c)). The adjustment includes (i) a constant offset, (ii) a reduced amplitude of the seasonal cycle and (iii) a phase-lag of one month. As shown in Fig. 2, after these adjustments the TRTS matches the mean seasonal cycle and absolute

value of the airborne data, thus accounting for known vertical gradients of CO_2 and reduced seasonality in the upper troposphere (Fig. 2, panel (d)). The amplitude of the seasonal cycle in our time series is also in good agreement with the boundary condition used by Andrews et al. (1999 and 2001).

The new TRTS is used to estimate the age spectrum in 13 N_2O bins of 5 or 10 ppb width (320-325 ppb,…, 290-295 ppb, 280-290,…, 230-240 ppb) using a Markov Chain Monte Carlo (MCMC) algorithm which compares observed CO_2 concentrations in the bin to concentration expected from the TRTS. To maximize data availability, we use high frequency data (see "2.1. Measurements") identified as stratospheric according to the criteria above from the 10 s merged products available for HIPPO, ORCAS, and ATom rather than data averaged to the lower Medusa sampling interval for this calibration exercise. Small corrections are applied to the observed CO_2 concentrations (<0.2 ppm) to account for oxidation of CH_4 and CO in the stratosphere.

The MCMC algorithm considers random noise in the TRTS and uses a unimodal or bimodal inverse Gaussian shape of the age spectrum characterized by mean ages Γ_1 and Γ_2 and shape parameters λ_1 and λ_2

$$G\left(t'|(\Gamma_1, \Gamma_2, \lambda_1, \lambda_2)\right)$$

$$= A\sqrt{\frac{\lambda_1}{2\pi t'^3}}\ \exp\left(-\frac{\lambda_1\ (t'-\Gamma_1)^2}{2\Gamma_1^2 t'}\right) + (1-A)\sqrt{\frac{\lambda_2}{2\pi t'^3}}\ \exp\left(-\frac{\lambda_2\ (t'-\Gamma_2)^2}{2\Gamma_2^2 t'}\right), \qquad 3$$

where factor A determines the relative weight of each peak and a value of $A = 1$ yields a unimodal spectrum (Hall and Plumb, 1994; Andrews et al., 2001; Bönisch et al., 2009). Though more complex representations of the age spectrum have recently been proposed, accounting for multiple modes and seasonal variations (e.g., Bönisch et al., 2009; Li et al., 2012; Hauck et al., 2019, 2020; Podglajen and Ploeger, 2019), these require additional information from multiple tracers or models, and good seasonal data coverage that is not available here. In any case, we only rely on the first moment of the age distribution (i.e., the mean age) for calibrating the N_2O-AoA relationship, and the mean age is not particularly sensitive to the assumed shape of the age spectrum (Andrews et al., 2001). By setting $A = 1$ for 50% of all tested age spectra, the MCMC algorithm automatically selects whether a unimodal or bimodal representation of the age spectrum is optimal to match observations. Although bimodal solutions with five instead of two free parameters will always be able to fit the data better, a larger number of parameters also decreases the likelihood of randomly selecting a combination of parameters that match the observations well because the fraction of the total parameter space region which yields good agreement with the observation decreases as more parameters are added (Malinverno, 2002). The algorithm is thus able to make an appropriate choice between unimodal and bimodal distributions from the data itself, without any further a priori assumptions. The MCMC algorithm is set up separately with 2000 independent chains for each N_2O bin to account for uncertainty in the TRTS and obtain best estimates of the mean AoA, i.e., the first moment of Eq. 3. To simplify the algorithm, possible values of $\Gamma_1, \Gamma_,\ \lambda_1$ and λ_2 are repeatedly sampled from the same N_2O bin-specific prior distributions. Details of the algorithm are presented in Appendix A.

Finally, the resulting relationship between mean AoA and the N$_2$O concentration of each bin is fit separately for each hemisphere by a quadratic polynomial and the polynomial is evaluated at the N$_2$O value of each Medusa sample to pair every observation of δ(Ar/N$_2$) with an AoA. Uncertainty in N$_2$O and the polynomial fits are propagated by a Monte Carlo scheme. Overall, our method estimates the most likely mean AoA for each Medusa sample and improves upon previous methods by providing a framework for the treatment of uncertainty resulting from (i) analytical error, (ii) uncertainty in the shape of the age spectrum, and (iii) uncertainty in the composition of source gas introduced into the stratosphere

2.3. TOMCAT/SLIMCAT model

TOMCAT/SLIMCAT (hereafter TOMCAT) is an offline 3-D chemical tracer transport model that has been used extensively for studies of stratospheric ozone depletion and circulation (e.g., Chipperfield, 2006; Chipperfield et al., 2017; Krol et al., 2018). For this study, TOMCAT was run over 31 years, from 1988 to 2018, with a timestep of 30 minutes at 2.8°×2.8° horizontal resolution forced by the ERA-Interim reanalysis (Dee et al., 2011) at 60 vertical hybrid sigma- pressure (σ–p) levels up to ~60–65 km. The first 20 years (i.e. before the first flask observation) were treated as spin-up. The TOMCAT AoA tracer is initialized at the surface and corrected to a value of AoA = 0 just below the tropical tropopause. Vertical motion was calculated from the divergence of the horizontal mass fluxes. Although this approach gives slightly younger stratospheric AoA than using isentropic levels and radiative heating rates, it allows a more detailed treatment of tropospheric transport (Chipperfield, 2006; Monge-Sanz et al., 2007). The model simulation was sampled at the times and locations of the Medusa flask observations to provide a direct comparison between the measurements and model.

Following the methodology of Belikov et al. (2019), we include an additional vertical flux term in the model representing the GS of gases in the atmosphere. The vertical flux f_i (mol m^{-2} s^{-1}) of tracer i due to molecular diffusion in Earth's gravitational field is given as (Banks and Kockarts, 1973; Belikov et al., 2019)

$$f_i = -D_i N C_i \left[\frac{1}{C_i} \frac{\partial C_i}{\partial z} + \left(\frac{1}{H_i} - \frac{1}{H_{air}} \right) + \alpha_i \frac{1}{T} \frac{\partial T}{\partial z} \right], \qquad\qquad 4$$

where D_i is the tracer-specific binary molecular diffusivity in air (m^2 s^{-1}), N is the number density of air (mol m^{-3}), C_i is the mixing ratio, $H_i = \frac{RT}{gM_i}$ is the tracer-specific atmospheric equilibrium scale height (m), R is the fundamental gas constant (J K^{-1} mol^{-1}), g is the gravitational constant (m s^{-2}), M_i is the tracer-specific atomic or molecular mass (kg mol^{-1}), α_i is the tracer-specific thermal diffusivity (m^2 s^{-1}), and T is temperature (K). The three terms in Eq. (4) represent molecular diffusion driven by (i) vertical gradients in the mole fraction of i, (ii) pressure gradients caused by gravity and described by the barometric law, and (iii) temperature gradients (left to right). We neglect the first and third terms in the brackets in Eq. (4), leaving only the gravitational settling term $\left(\frac{1}{H_i} - \frac{1}{H_{air}} \right)$ on the basis that both terms are orders of magnitude less important than the gravitational

separation term under stratospheric conditions (see Appendix B) (Ishidoya et al., 2013; Belikov et al., 2019). No fluxes are allowed through the top boundary, and C_i is held constant at the Earth's surface.

210   To simplify the numerical treatment, we simulate an idealized reference tracer of gravitational fractionation δ_{GST}, which has a molecular mass 1 amu greater than air and diffusivity equal that of Ar(see Appendix B & C). This tracer can be scaled offline to obtain the gravitational separation signal in any other species, including δ(Ar/N₂), e.g.

$$\delta(Ar/N_2) \approx \frac{(M_{Ar} - M_{air}) \times D_{Ar} - (M_{N2} - M_{air}) \times D_{N2}}{(M_{GST} - M_{air}) \times D_{GST}} \times \delta_{GST}. \qquad\qquad 5$$

The appropriate diffusivity values D_{Ar} and D_{N2} for Ar and N_2 in air are derived in Appendix B for a ternary mixture of Ar, O₂, and N₂, extending previous work (Ishidoya et al., 2013; Belikov et al., 2019). δ_{GST}.is divided by D_{GST} in eq. 5 and therefore

215   does not depend on the exact value chosen for D_{GST}.

2.4. NIES TM model

We compare our results to previous simulations of GS using the National Institute for Environmental Studies chemical transport model (NIES TM) published recently by Belikov et al. (2019). The NIES TM is a three-dimensional transport model of similar complexity as TOMCAT driven by the Japanese 25-year Reanalysis (JRA-25) with a hybrid sigma–isentropic (σ–

220   θ) vertical coordinate up to 5 hPa or ~35 km. The model and GS results are described in detail in Belikov et al. (2013) and Belikov et al. (2019), respectively.

3. Results

3.1. N₂O-AoA calibration

Our new relationships between N₂O and mean AoA for the Northern and Southern Hemispheres (NH & SH) are well

225   constrained by the observations and generally follow the mid-latitude NH calibration curve of Andrews et al. (2001) (Fig. 3, panel (d)). At mean AoA>2.5, the N₂O-AoA relationship yields slightly younger ages in the NH than in the SH, and compared to the previously published curve, suggesting there might be a latitudinal dependence of the relationship. Such a latitude-dependence is, in fact, expected based on theory (Plumb, 2007). We expect curvature and a latitudinal difference in the N₂O-AoA relationship because photolysis of N₂O depends on latitude and altitude due to local sunlight availability. Furthermore,

230   mixing of young and old air results in a mixture with anomalously low N₂O concentration for a given age. Since the SH, NH and tropics feature different photolysis rates and show different degrees of mixing/isolation, different N₂O-AoA relationship are expected.

Unimodal age spectra are generally preferentially selected by the algorithm, in particular for the SH (Fig. 3, panels (a)-(c)) but unimodal age spectra dominate the solution ensemble by a small margin except for AoA <0.5 years (Fig. S3). Confidence

intervals on age spectra parameters from bimodal spectra are considerably wider than for unimodal spectra. This implies that the parameters in a bimodal distribution are redundant and the shape of the age spectrum is not sufficiently constrained by the observations used. The width of the spectrum in each N_2O bin varies widely within the prescribed bounds for most unimodal and bimodal age spectra (Fig. S3). It appears that not enough data is available from the airborne campaigns to determine the amplitude of the seasonal cycle with enough confidence to constrain the width of the age spectrum and distinguish the relative contribution of the old peak (influencing primarily the mean concentration difference to the troposphere) and the young peak (controlling the amplitude of seasonality) in bimodal age spectra. Additional observations with less scatter or information from different age tracers are needed to properly resolve the shape of the age spectra. Despite the limited resolution of the seasonal cycle, the observations are sufficient to place tight limits on the mean AoA in each N_2O bin and yield a well-characterised relationship between N_2O and mean AoA for each hemisphere.

3.2. The relationship between mean AoA and GS in models and observations

A comparison of the AoA-GS relationship with observations yields good agreement for the TOMCAT model results within the observational uncertainties (Fig. 4 panel (a)), but the observations fall outside the range of GS predicted by the NIES TM (Belikov et al., 2019) for mean AoA >1 year (Fig. 4, panel (b)). For young samples with mean AoA < 3 years, GS of $\delta(Ar/N_2)$ increases by roughly 35-45 per meg per year of AoA in both TOMCAT and observations, and converges to zero for the youngest samples. In the upper stratosphere, TOMCAT does not obtain any ages as old as observed by Ishidoya et al., (2008, 2013, 2018) and therefore cannot reproduce these observations directly. Changing the vertical coordinate system of TOMCAT or forcing the model with a different reanalysis product could improve agreement with the observations for old ages because TOMCAT in the configuration used here is known to slightly underestimate AoA in the upper stratosphere (Monge-Sanz et al., 2007; Chabrillat et al., 2018). The very steep AoA-GS relationship for the oldest simulated air is however seen in TOMCAT and the balloon observations.

A comparison of $\delta(Ar/N_2)$ and AoA between flask samples and TOMCAT is shown in Figure S4. However, the direct comparison should not be considered a good test of model performance because our stratospheric composition criteria intentionally target unusually old air in the lowermost stratosphere. For the same reason, profiles of $\delta(Ar/N_2)$ and AoA are not evaluated here. Instead, we suggest the AoA-GS relationship provides a more robust metric for comparison. Overall, our observations validate the implementation of GS for young (< 3 y) ages in TOMCAT.

The relationship between mean AoA and GS differs between TOMCAT and the NIES TM (Fig. 5). The NIES TM shows weaker curvature in the relationship overall and produces larger declines in Ar/N_2 at ages less than 4.5 years compared to TOMCAT. In TOMCAT's mesosphere (which is at the limit of the domain covered by the ERA-Interim reanalyses forcing the model), AoA is near uniform but GS continues to increase with increasing altitude, changing the relationship between mean AoA and GS in this region. The AoA-GS relationship in TOMCAT is very similar in all non-tropical regions (outside 15° > lat > -15°) whereas the curvature of the relationship is slightly stronger in the tropics. In contrast, the NIES TM has a

clear dependence of the AoA-GS relationship on latitude. There is also some evidence in the observations for a dependence of the AoA-GS relationship on hemisphere. In the observations, $\delta(Ar/N_2)$ values appear to be slightly more negative in the SH than in the NH for the same age, particularly for AoA > 1.5 years. However, almost all SH samples with mean AoA older than 1.5 years come only from the ORCAS campaign. This campaigned used a different inlet design than HIPPO and ATom which could cause a small artifactual offset but the scatter in our observations generally makes it difficult to separate signal from noise for small interhemispheric differences.

3.3. GS and AoA in TOMCAT

Annual mean $\delta(Ar/N_2)$ in TOMCAT follows the typical pattern of a tracer with a stratospheric sink (Chipperfield, 2006), as previously found in simulations using the NIES TM (Belikov et al., 2019) and the SOCRATES model (Ishidoya et al., 2013; Sugawara et al., 2018) (Fig. 6). $\delta(Ar/N_2)$ is zero in the troposphere and decreases with elevation. The most depleted $\delta(Ar/N_2)$ is observed at high latitudes where sinking air of the Brewer-Dobson circulation advects strongly fractionated air downward. In the tropics, $\delta(Ar/N_2)$ values are considerably less fractionated at the same altitude due to upwelling of unfractionated tropospheric air. Vertical gradients in $\delta(Ar/N_2)$ generally increase with altitude and are largest in the mesosphere (> 50km) because molecular diffusion increases with decreasing pressure and eventually dominates above the turbopause (not shown). There are strong seasonal changes in $\delta(Ar/N_2)$ depletion on the order of several thousand per meg, in particular in the high latitude mesosphere, with the strongest fractionation occurring during the winter season (see movie available with online version of the manuscript).

The AoA tracer in TOMCAT shows a similar pattern to $\delta(Ar/N_2)$ with younger ages at low altitude and in the tropics and oldest ages in the mesosphere. Vertical gradients in mean AoA are largest at high latitudes close to the tropopause. At high latitudes above 20 km and at low latitudes above 50 km, vertical gradients of AoA mostly disappear and AoA becomes nearly uniform.

4. Discussion

4.1. Difference between TOMCAT and the NIES TM

We hypothesize that an adequate representation of the mesosphere in models is critical in determining the curvature of the AoA-GS relationship. The residence time of air above ~40 km is rather short in TOMCAT and AoA is nearly constant with altitude in this region. GS in contrast continues to increase with altitude because of the diffusivity dependence on pressure allowing gases to separate more effectively. The much lower top of the NIES TM (~35 km vs 60–65 km) reduces its ability to capture this effect, which impacts the AoA-GS relationship because the mesospheric signal is exported into the stratosphere, in particular in polar regions where mesospheric air descends. TOMCAT furthermore produces a less negative slope of the AoA-GS relationship for young air (mean AoA <3 years) and greater similarity in the AoA-GS relationship between latitude

bands than the NIES TM as shown in Fig. 5. This could in part be a consequence of using a different meteorological reanalysis product for forcing the two models (Chabrillat et al., 2018), or could indicate differences between the models in vertical and horizontal mixing intensity.

4.2. Estimating mean AoA from observed δ(Ar/N₂)

Different tracers of AoA all have unique strengths and weaknesses. Estimating AoA in the lowermost stratosphere from CO_2 for example is limited by complexities involving seasonality and the possibility of multiple entry points of air into the stratosphere due to isentropic mixing with the midlatitude troposphere (Hall and Plumb, 1994; Andrews et al., 2001; Waugh and Hall, 2002). SF_6-derived AoA in contrast is biased high at mid and high latitudes due to the influence of mesospheric air which has been photochemically depleted in SF_6 (Kovács et al., 2017; Linz et al., 2017). N_2O as an AoA tracer relies on the photolytic destruction of N_2O in the stratosphere which may depend strongly on location and had only been empirically calibrated for young ages at mid latitudes so far (Andrews et al., 2001; Linz et al., 2017).

Thanks to the robust relationship between mean AoA and δ(Ar/N₂), and the small seasonal cycle amplitude of δ(Ar/N₂) <5 per meg in the upper troposphere (Fig. S1), mean AoA could also be estimated from δ(Ar/N₂). Using the current analytical δ(Ar/N₂) uncertainty of 12.2 per meg (2σ) and the AoA-δ(Ar/N₂) relationship seen in TOMCAT (including variability from seasonal and latitudinal differences), we estimate that mean AoA could be calculated to within about ±0.4 years (2σ, Fig. S5). This is still considerably worse than the ±0.1 years confidence interval in mean AoA estimated from N_2O. However, the uncertainty can be improved with future improvements in sampling, and the accuracy and precision of the measurements.

The heavy noble gases krypton and xenon will be roughly 3.6× and 5.8× more strongly fractionated in the stratosphere than Ar, but are also more challenging to measure due to their ~8000× and 100000× lower abundance in the atmosphere. Nevertheless, future analysis of these gases in stratospheric air samples might further improve our ability to estimate AoA from the gravitational fractionation of gases and help diagnose artefactual fractionation, because heavier noble gases are more strongly fractionated under the influence of gravity and less sensitive to thermal fractionation (Seltzer et al., 2019).

4.3. Future directions

An open question in climate applications of noble gases is whether there could be a stratospheric influence on tropospheric δ(Ar/N₂). Tropospheric observations of δ(Ar/N₂) and other noble gas-elemental ratios have been used to infer ocean heat content changes by capitalizing on the temperature-dependency of gas solubility in the oceans (Keeling et al., 2004; Headly and Severinghaus, 2007; Ritz et al., 2011; Bereiter et al., 2018). However, long-term trends (Butchart, 2014) and natural variability in stratospheric circulation and stratosphere-troposphere exchange (STE) such as the Quasi-Biennial Oscillation (QBO) (Baldwin et al., 2001) could advect a stratospheric GS signal into the troposphere and alias onto surface observations of these gases. To this end, we calculate the volumetric Ar deficit in the atmosphere in moles using TOMCAT (Fig. 7) relative to a hypothetical, unfractionated atmosphere with a homogenous mixing ratio ($C_{o_{Ar}}$) at all elevations. Ar deficit is

$$Ar_{deficit} = \left(C_{Ar} - C_{o_{Ar}}\right) \times N_{air} \qquad\qquad 6$$

where C_{Ar} is the simulated Ar mixing ratio and N_{air} is the number density of air.

The Ar deficit is concentrated in the lower stratosphere at around 20 km and at mid to high latitudes. Although the GS signal
330   is considerably stronger in the mesosphere, the potential for perturbing the troposphere is low given the low molar density of
air in the mesosphere. The region with greatest potential to influence the troposphere therefore lies in the lower stratosphere.
The total Ar deficit of the atmosphere above 200 hPa is approximately -2.9 $\times 10^{13}$ mol and the atmosphere below 200 hPa
contains roughly 1.3 $\times 10^{18}$ mol of Ar. Perturbing STE and/or the stratospheric circulation by 10%, consistent with interannual
to decadal variability of STE in models (Salby and Callaghan, 2006; Flury et al., 2013; Ray et al., 2014; Montzka et al., 2018),
335   thus should lead to a detectable signal of roughly 2.2 per meg (-2.9$\times 10^{13}$/1.3$\times 10^{18}\times$10%$\times 10^6$) in tropospheric δ(Ar/N$_2$).
Corresponding advection of stratospheric GS signals in N$_2$ amplifies the pure Ar signal by roughly 10%. A careful investigation
of such a signal in the δ(Ar/N$_2$) surface network data (Keeling et al., 2004) is needed because secular trends of δ(Ar/N$_2$) caused
by degassing of Ar and N$_2$ from a warming ocean are also expected to be on the order of 2-3 per meg per decade. Previous
studies have also used ratios involving heavier noble gases (Xe/N$_2$, Kr/N$_2$) to reconstruct mean ocean temperature changes
340   over glacial-interglacial timescales (Headly and Severinghaus, 2007; Bereiter et al., 2018; Baggenstos et al., 2019).
Simultaneous changes in stratospheric circulation and STE affect heavier noble gases more strongly than Ar/N$_2$ and will need
to be accounted for in such applications of noble gas thermometry.

5.   Conclusions

With improvements in data treatment, measurement quality, and modelling constraints, we have shown that gravitational
345   fractionation of Ar relative to N$_2$ in the stratosphere and mesosphere is a potentially powerful constraint on circulation. High-
precision observations of δ(Ar/N$_2$) in air samples of the lowermost stratosphere from 9 airborne campaigns are well captured
by the 3-D chemical transport model TOMCAT/SLIMCAT, which has been updated to account for the influence of gravity on
air composition through molecular diffusion. In the observations and the model, δ(Ar/N$_2$) is directly related to stratospheric
mean age of air (AoA) derived here using a new calibration of N$_2$O. Our observations for mean AoA <3 years produce a slope
350   of roughly 35-45 per meg of δ(Ar/N$_2$) per year of AoA. TOMCAT/SLIMCAT shows better agreement with the new
observations than the NIES transport model (Belikov et al., 2019) and we speculate that the model disagreement could be
explained by (i) the factor of 2 lower top of the NIES transport model, (ii) the use of different reanalysis products, and/or (iii)
differences in vertical and horizontal mixing. In this context, further work is also needed to study the influence of unresolved
turbulence on AoA and δ(Ar/N$_2$) in chemical transport models.

355   As the importance of stratospheric circulation for ozone recovery, climate projections, and evaluation of tropospheric trends
in halocarbons is increasingly recognized, a need for new observations from the still undersampled stratosphere is becoming

evident. Combining $\delta(Ar/N_2)$ with other tracers of circulation could lead to new insights into atmospheric mixing and transport. $\delta(Ar/N_2)$ has potential advantages over existing approaches based on transient tracers such as CO_2 or N_2O since $\delta(Ar/N_2)$ is only influenced by the physics of transport and mostly unaffected by seasonality in the tropical upper troposphere. Furthermore,

360 because of its large gradients at high altitudes, $\delta(Ar/N_2)$ observations from the upper stratosphere and mesosphere could improve our understanding of circulation on seasonal and interannual timescales in a region where AoA becomes near-uniform and some photochemical age tracers such as N_2O are almost fully depleted and thus lose sensitivity to circulation changes.

6. Data availability

All Data from the ORCAS, HIPPO and ATom airborne campaigns are freely available at: doi:10.5065/D6SB445X,

365 doi:10.3334/CDIAC/HIPPO_014, and doi:10.3334/ornldaac/1581. All of the primary data used here is consistent with the associated merged data products in Wofsy et al. (2012), Stephens (2017), and Wofsy et al. (2018). For Medusa samples identified as stratospheric, CO_2, $\delta(Ar/N_2)$, AoA, and kernel-averaged N_2O, O_3, H_2O, CO, CH_4, as well as additional metadata are available as a supplement to this manuscript at https://doi.org/10.3334/ORNLDAAC/1788.

NIES TM modelling results (Belikov et al., 2019) and stratospheric observations of GS from the cryogenic balloon sampling-

370 system (Ishidoya et al., 2008, 2013, 2018; Sugawara et al., 2018) were provided directly by the authors.

7. Appendix A: Description of the Markov chain Monte Carlo algorithm

The following list outlines key steps in our MCMC algorithm to calculate AoA from CO_2 for each N_2O bin:

1) Start a new Markov Chain and allow for uncertainty in TRTS by adding white noise with an amplitude given by the scatter of upper tropospheric observations around the mean TRTS in Fig. 2.

375 2) With a 50% chance, peak weighting factor A is set to zero and a unimodal spectrum is tested (k=1). Alternatively, A is allowed to vary between 0 and 1 for a bimodal distribution (k=2).

3) The other age spectrum parameters are selected from N_2O bin-dependent prior distributions: Γ_1 is sampled from a uniform prior distribution with a mean AoA predicted by the N_2O-AoA relationship of Andrews et al. (2001) and with generous width (>1.5 years). Γ_2 is sampled from a uniform prior distribution with values between 1-6 years

380 representing older stratospheric air. The shape parameters are defined as $\lambda_i \equiv \frac{\Gamma_i^2}{2\,\gamma_i}$, and γ_i is chosen randomly for each peak with values between 0.1 and 1, as observed in previous studies (Hall and Plumb, 1994; Andrews et al., 2001; Waugh and Hall, 2002).

4) Convolve the age spectrum calculated from the parameters ($\boldsymbol{m} \equiv [\Gamma_1, \Gamma_2, \lambda_1, \lambda_2, A]$) with the perturbed TRTS to obtain a possible time series of CO_2 in the N_2O bin and calculate the misfit ($\boldsymbol{e}$) between the CO_2 time series and observations.

385   5) Calculate the likelihood function $P(\boldsymbol{d}|k,\boldsymbol{m})$ for the set of parameters $\boldsymbol{m}$ and k, given a total of n observations $(\boldsymbol{d})$

$$P(\boldsymbol{d}|k,\boldsymbol{m}) = \frac{1}{[(2\pi)^n \det(\widehat{\boldsymbol{C}}_e)]^{0.5}} \exp\left(-\frac{1}{2}e^T \widehat{\boldsymbol{C}}_e^{-1} e\right) \tag{A1}$$

where $\widehat{\boldsymbol{C}}_e$ is the covariance matrix. We assume that all n observations are independent. Thus, $\widehat{\boldsymbol{C}}_e$ has only diagonal entries of $\sigma_{CO_2}^2$ and $\det(\widehat{\boldsymbol{C}}_e)$ simplifies to $\left(\sigma_{CO_2}^2\right)^n$. The value of $\sigma_{CO_2}^2$ is different for each bin and determined iteratively as the approximate root mean square error of the observations around the final time series for each bin obtained at the end of the MCMC algorithm. Typical values of $\sigma_{CO_2}^2$ are between 0.18 and 1.28 ppm and generally

390   decrease with increasing AoA of a N_2O bin.

6) If this is the first pass of the chain, define k_{saved} and $\boldsymbol{m}_{saved}$ to equal k and $\boldsymbol{m}$. Otherwise, calculate the selection criterion $\alpha \equiv \min\left(1, \frac{P(\boldsymbol{d}|k,\boldsymbol{m})}{P(\boldsymbol{d}|k_{saved},\boldsymbol{m}_{saved})}\right)$ and accept k and $\boldsymbol{m}$ as new saved values $(k_{saved}, \boldsymbol{m}_{saved})$ with probability α. Sampling from the same prior distributions on each pass of the chain simplifies our expression of α compared to that presented by Malinverno (2002), making it only dependent on the likelihood ratio.

395   7) Repeat steps 2-6 1000 times sampling parameter values from the same prior distributions and store the final values of k_{saved} and $\boldsymbol{m}_{saved}$ obtained after the 1000th iteration (i.e., a plausible solution produced past the burn-in period) for later use.

8) To sample the full posterior pdf (i.e., the full uncertainty about the age spectrum parameters), initialize 2000 different Markov chains by repeating steps 1-7. Each stored value of $\boldsymbol{m}$ characterizes one age spectrum that is likely not far

400   from the best solution, given the data $\boldsymbol{d}$, yielding an ensemble of 2000 age spectra from which statistics can be computed. Note that each Markov chain is fully independent, so the algorithm can be easily parallelized to minimize computational costs.

8. Appendix B: Derivation of Eq. (4) from the Maxwell-Stefan Equations

We start by approximating air as a ternary mixture of N_2, O_2 and Ar, and later generalize to consider additional trace species.

405   According the Maxwell-Stefan equations (Taylor and Krishna, 1993) diffusion in this ternary mixture is governed by:

$$d_{N2} = \frac{C_{N2}\, f_{Ar} - C_{Ar}\, f_{N2}}{N \times D_{N2:Ar}} + \frac{C_{N2}\, f_{O2} - C_{O2}\, f_{N2}}{N \times D_{N2:O2}} \tag{B1}$$

$$d_{Ar} = \frac{C_{Ar}\, f_{N2} - C_{N2}\, f_{Ar}}{N \times D_{Ar:N2}} + \frac{C_{Ar}\, f_{O2} - C_{O2}\, f_{Ar}}{N \times D_{Ar:O2}} \tag{B2}$$

$$f_{O2} = -f_{N2} - f_{Ar} \tag{B3}$$

$$C_{N2} + C_{O2} + C_{Ar} = 1 \tag{B4}$$

where $C_i \equiv n_i/N$ is the mole fraction, n_i is molar or number density (mol m^{-3}), N is the total number density (mol m^{-3}), f_i is the molecular diffusion flux (mol m^2 s^{-1}) relative to the molar average velocity of the mixture, d_i is the thermodynamic driving force for molecular diffusion (m^{-1}), and $D_{i:j}$ is the binary diffusion coefficient of the $(i:j)$ pair. An equation for O$_2$ that is analogous to Eq. (B1) is not needed because changes in O$_2$ are governed by the conservation Eqs. (B3) and (B4). Binary
410   diffusivity coefficients (cm^2 s^{-1}) can be calculated using the method of Fuller et al. as reported in (Reid et al., 1987)

$$D_{ij} = 0.001 \frac{T^{1.75} \left(\frac{1}{M_i} + \frac{1}{M_j} \right)^{0.5}}{P \left[v_i^{1/3} + v_j^{1/3} \right]^2} \tag{B5}$$

where P is pressure (atm), v_i is the molecular diffusion volume of a trace gas or air (Table B1), and M_i is the molecular mass (g mol^{-1}) of a gas.

For an ideal gas, d_i is given by

$$d_i \equiv \nabla C_i + (C_i - \omega_i)\frac{\nabla P}{P} - \frac{\rho_i}{P}\left(\frac{F_i}{M_i} - \sum_{j=1}^{n} \omega_i \frac{F_j}{M_j} \right) + \frac{k_i^T}{T}\nabla T \tag{B6}$$

where $\omega_i \equiv \frac{\rho_i}{\rho} = \frac{M_i C_i}{M_{air}}$ is the mass fraction of gas i, ρ_i is density of i (kg m^{-3}), P is pressure (Pa), T is temperature (K), k_i^T is
415   the thermal-diffusion ratio of i, M_i is the molecular mass of i (kg mol^{-1}), and F_i the external body force per mole (N mol^{-1}) for i (Chapman et al., 1990; Taylor and Krishna, 1993).

In the atmosphere, (vertical) pressure gradients are caused by gravity and well approximated by hydrostatic balance

$$\nabla P \approx \frac{\partial P}{\partial z} \approx -\rho g = -\frac{P M_{air}}{RT} g \tag{B7}$$

The gravitational force per mole F_i is

$$F_i = -\frac{\rho_i}{n_i} g \tag{B8}$$

Substituting Eqs. (B7) and (B8) into Eq. (B6) yields

$$d_i \approx \nabla C_i - \left(1 - \frac{M_i}{M_{air}}\right) C_i \frac{1}{P} \frac{P M_{air} g}{RT} + \frac{\rho_i}{P}\left(\frac{\frac{\rho_i}{n_i} g}{M_i} - \sum_{j=1}^{n} \omega_j \frac{\frac{\rho_j}{n_j} g}{M_j}\right) + \frac{k_i^T}{T} \nabla T$$

$$= \nabla C_i + (M_i - M_{air})\frac{g}{RT} C_i + \frac{\rho_i}{P}\left(g - \sum_{j=1}^{n} \omega_i g\right) + \frac{k_i^T}{T}\nabla T \tag{B9}$$

$$= \nabla C_i + \left(\frac{1}{H_i} - \frac{1}{H_{air}}\right) C_i + \frac{k_i^T}{T}\nabla T$$

420   where we use the definition of the scale height $H_i \equiv \frac{RT}{g M_i}$. The two terms involving the body force cancel because all species experience the same gravitational force per unit mass. The tendency for gravimetric separation instead arises from the pressure gradient term proportional to $\left(\frac{1}{H_i} - \frac{1}{H_{air}}\right)$.

Equations (B1) and (B2) can be inverted to solve for f_{N2} and f_{Ar} (Taylor and Krishna, 1993)

$$f_{N2} = -N D_{N2}^{air} d_{N2} - N D_{N2\times(Ar,O2)}^{air} d_{Ar} \tag{B10}$$

$$f_{Ar} = -N D_{Ar\times(N2,O2)}^{air} d_{N2} - N D_{Ar}^{air} d_{Ar} \tag{B11}$$

where

$$D_{N2}^{air} = \frac{D_{N2:O2}(C_{N2} D_{Ar:O2} + (1 - C_{N2}) D_{N2:Ar})}{S} \tag{B12}$$

$$D_{N2\times(Ar,O2)}^{air} = \frac{C_{N2} D_{Ar:O2}(D_{N2:O2} - D_{N2:Ar})}{S} \tag{B13}$$

$$D_{Ar}^{air} = \frac{D_{Ar:O2}(C_{Ar} D_{N2:O2} + (1 - C_{Ar}) D_{N2:Ar})}{S} \tag{B14}$$

$$D_{Ar\times(N2,O2)}^{air} = \frac{C_{Ar} D_{N2:O2}(D_{Ar:O2} - D_{N2:Ar})}{S} \tag{B15}$$

$$S = C_{N2} D_{Ar:O2} + C_{Ar} D_{N2:O2} + C_{O2} D_{N2:Ar} \tag{B16}$$

425   Here D_{N2}^{air} and D_{Ar}^{air} are the effective diffusivities of N_2 and Ar in air, while $D_{N2\times(Ar,O2)}^{air}$ and $D_{Ar\times(N2,O2)}^{air}$ reflect ternary cross interactions, such as the tendency of N_2 to be impacted by any process that drives a diffusive flux of Ar.

In air, Ar is a minor gas ($C_{Ar} \ll C_{N2} \sim C_{O2}$) and therefore interactions of N_2 with Ar can be neglected in the N_2 flux and the diffusive fluxes of N_2 and O_2 must balance approximately as in the case of a binary mixture of the two gases

$$f_{N2} \approx -N D_{N2}^{air} d_{N2} \tag{B17}$$

$$f_{O2} \approx -f_{N2} \tag{B18}$$

Combining Eqs. (B10) and (B17) with Eq. (B9), yields

$$f_{N2} \approx -N D_{N2}^{air} \left[\nabla C_{N2} + \left(\frac{1}{H_{N2}} - \frac{1}{H_{air}} \right) C_{N2} + \alpha_{N2:O2}^{T} \frac{C_{N2}(1 - C_{N2})}{T} \nabla T \right] \tag{B19}$$

$$f_{Ar} \approx -N D_{Ar \times (N2,O2)}^{air} \left[\nabla C_{N2} + \left(\frac{1}{H_{N2}} - \frac{1}{H_{air}} \right) C_{N2} + \alpha_{N2:O2}^{T} \frac{C_{N2}(1 - C_{N2})}{T} \nabla T \right]$$

$$- N D_{Ar}^{air} \left[\nabla C_{Ar} + \left(\frac{1}{H_{Ar}} - \frac{1}{H_{air}} \right) C_{Ar} + \alpha_{Ar:air}^{T} \frac{C_{Ar}}{T} \nabla T \right] \tag{B20}$$

430   where we have replaced the thermal diffusion ratio k_i^T with the better empirically constrained thermal diffusion factor $\alpha_i^T \equiv \frac{k_i^T}{C_i C_j}$ defined as such only in binary mixtures. Therefore, $k_{N2}^T \approx \alpha_{N2:O2}^T C_{N2} C_{O2}$ and $k_{Ar}^T \approx \alpha_{Ar:air}^T C_{Ar}$. Table B2 presents rough estimates of the magnitudes of the terms in Eqs. (B19) and (B20) in the stratosphere, showing that the thermal diffusion and cross-diffusion terms involving $D_{Ar \times (N2,O2)}^{air}$ are at least two orders of magnitude smaller than the remaining terms. Neglecting these smaller terms yields the governing Eq. (4) used in our model simulation

$$f_i \approx -N D_i \left[\left(\frac{1}{H_i} - \frac{1}{H_{air}} \right) C_i \right] = -N D_i \Delta M_i \frac{g}{RT} C_i \tag{B21}$$

435   where we introduced $\Delta M_i \equiv M_i - M_{air}$ the molecular mass difference to air (kg mol^{-1}) as a short-hand. Equation (B21) is equally valid for trace gases such as Ar and major gases N_2 and O_2 when the appropriate diffusivities given by Eqs. (B12) and (B14) are used. In our case $D_{N2} \equiv D_{N2}^{air} = D_{O2}^{air} \approx D_{N2:O2}$ for N_2 and $D_{Ar} \equiv D_{Ar}^{air} \approx \frac{D_{Ar:O2} D_{N2:Ar}}{C_{N2} D_{Ar:O2} + C_{O2} D_{N2:Ar}}$ for Ar.

9.   Appendix C: Calculating δ(Ar/N₂) from δ(GST)

The conservation equation of gas i with mixing ratio C_i accounting for advection (1st term RHS), eddy mixing (2nd term RHS),

440   and molecular diffusion (3rd term RHS, using the simplified Eq. (B21)) is given by

$$\frac{\partial}{\partial t}(C_i) = -\vec{u} \cdot \vec{\nabla}[C_i] + \vec{\nabla} \cdot \left[\boldsymbol{D}_e \vec{\nabla} C_i \right] + \frac{\partial}{\partial z} \left[D_i \Delta M_i \frac{g}{RT} C_i \right] \tag{C1}$$

where $\vec{u}$ is the velocity vector (m s^{-1}), $\boldsymbol{D}_e$ is the eddy diffusivity tensor (m^2 s^{-1}), D_i is the molecular diffusivity of species i in air (i.e., D_i^{air} m^2 s^{-1}), g the gravitational acceleration, R is the fundamental gas constant (J K^{-1} mol^{-1}), and T is temperature (K). "$\vec{\nabla}\cdot$" represents the divergence operator. $\vec{u}, \boldsymbol{D}_e, D_i, N$, and T depend on x,y,z, and t and the largest gradients for these variables generally occur in the vertical direction.

445   Dividing Eq. (C1) by a reference value $C_{i,0}$, separating C_i into a constant and a perturbation component (i.e., $\frac{C_i}{C_{i,0}} = 1 + \frac{C_i'}{C_{i,0}} = 1 + \delta_i$), and using the chain rule yields

$$\frac{\partial}{\partial t}(\delta_i) = -\vec{u}\cdot\vec{\nabla}[\delta_i] + \vec{\nabla}\cdot\left[\boldsymbol{D}_e\vec{\nabla}\delta_i\right] + D_i\Delta M_i\frac{g}{RT}\frac{\partial}{\partial z}[\delta_i] + (1+\delta_i)\frac{\partial}{\partial z}\left[D_i\Delta M_i\frac{g}{RT}\right] \qquad (C2)$$

To simplify Eq. (C2), we assume that the perturbations in the mixing ratio are small (i.e., $\delta_i \ll 1$) which is true at the 1% level for $\delta(Ar/N_2)$ even in the mesosphere. Furthermore, we assume that $U + \frac{D_e}{L} \gg \frac{D_i}{\Delta H_i}$ (U is a characteristic velocity scale, D_e is a characteristic eddy diffusivity, and L is a characteristic length scale; $\Delta H_i = \frac{RT}{\Delta M_i g}$) or equivalently stated in terms of the Peclet

450   number $Pe \sim \frac{U\Delta H_i + D_e\frac{\Delta H_i}{L}}{D_i} \gg 1$. We estimate typical values of the vertical Peclet number in the stratosphere to be between ~1000 to 10000 based on the height (~20 km) and turnover time (~5 years) of the stratosphere and the range of molecular diffusivities given in Table B1. In the mesosphere, AoA is near-uniform and we conservatively assume a turnover time of 0.5 years and height of 20 km. This implies a Peclet number of ~500 for the mesosphere. Under these conditions, the third term on the RHS can be eliminated because it is $O(Pe)$ smaller than the first and second terms and because it is $O(\delta_i)$ smaller than the fourth

455   term. Thus, we obtain

$$\frac{\partial}{\partial t}(\delta_i) \approx -\vec{u}\cdot\vec{\nabla}[\delta_i] + \vec{\nabla}\cdot\left[\boldsymbol{D}_e\vec{\nabla}\delta_i\right] + \frac{\partial}{\partial z}\left[D_i\Delta M_i\frac{g}{RT}\right] \qquad (C3)$$

Where the last term approximates the vertical divergence of the gravitational flux. At the top of the atmosphere and Earth's surface, the gravitational flux abruptly vanishes, and its divergence becomes large. If its divergence is small in the interior, Eq. (C3) can be conceptually interpreted as an advection-diffusion problem with large apparent sources or sinks at the bottom and top boundary due to diffusive flux divergence. The steady-state solution to Eq. (C3) for a tracer δ_i with no additional sources

460   and sinks (other than the apparent sources from diffusive flux divergence) yields the steady-state profile for species i. This solution can be scaled to yield the steady-state solution for any other tracer δ_j as

$$\delta_j \approx \frac{D_j\Delta M_j}{D_i\Delta M_i}\delta_i \qquad (C4)$$

as can easily be validated by solving Eq. (C4) for δ_i and substitution into Eq. (C3). Note that $\frac{D_j\Delta M_j}{D_i\Delta M_i}$ does not depend on x,y, and z.

Furthermore, ratios of passive tracers k and l can be calculated directly from δ_i by recognizing that $\delta_{k/l} \approx \delta_k - \delta_l$ for $\delta_k + \delta_l \ll 1$. Hence, we obtain the general version of Eq. (5) in the main text

$$\delta_{k/l} \approx \delta_k - \delta_l \approx \frac{D_k^{air}\Delta M_k - D_l^{air}\Delta M_l}{D_i \Delta M_i}\delta_i. \qquad (C5)$$

10. Author contributions

MPC, WF and CW set up and ran the TOMCAT model simulations. EJM, JDB, BBS, and SCW collected samples and curated data. BB carried out the data analysis with important input from MPC, RFK, SCW, ML, and EJM. BB prepared the manuscript with contributions from all co-authors.

11. Competing Interests

The authors declare that they have no conflict of interest.

12. Acknowledgements

We would like to thank the pilots and crew of the GV and DC-8 research aircraft as well as NCAR and NASA project managers, field support staff, and logistics experts that were integral to the success of the ORCAS, HIPPO and ATom campaigns. We are grateful to Andy Watt at NCAR who helped prepared the Medusa sampler for the field and assisted with its operation. Bill Paplawsky, Sara Afshar, Shane Clark, Stephen Walker and Adam Cox facilitated sample analysis, gas cylinder preparation, data management and performed instrument maintenance at Scripps Institution of Oceanography (SIO). CO_2 flask analysis at SIO is supported by NASA grant NNX17AE74G. Special thanks go to Eric Hintsa, Fred Moore, Jim Elkins, Mark Zondlo, Stuart Beaton, Minghui Diao, Glenn Diskin, Glen Sachse, Joshua DiGangi, John Nowak, James Flynn, Sergio Alvarez, Kathryn McKain, and Colm Sweeney for sharing O_3, N_2O, H_2O, CO and CH_4 data used in this manuscript.

For VCSEL hygrometer, field support, laboratory calibrations and QA/QC were provided by M. Diao, J. DiGangi, M. Zondlo and S. Beaton in HIPPO; S. Beaton provided field support and M. Diao provided laboratory calibrations and QA/QC in ORCAS.

This material is based upon work supported by the National Center for Atmospheric Research (NCAR), which is a major facility sponsored by the National Science Foundation (NSF) under Cooperative Agreement No. 1852977. The HIPPO program was supported by NSF grants ATM-0628575, ATM-0628519 and ATM-0628388 to Harvard University, University of California (San Diego), NCAR, and University of Colorado/CIRES. The primary ORCAS program was supported by NSF Polar Programs Grants 1501993, 1501997, 1501292, 1502301, and 1543457 to NCAR, SIO, University of Colorado/CIRES,

and the University of Michigan. The ATom program was supported by the NASA grant NNX15AJ23G. Medusa sampling on

490 ATom was supported by NSF grants AGS-1547797 and AGS-1623748 to University of California (San Diego) and NCAR. The TOMCAT simulations were performed on the UK Archer and University of Leeds ARC HPC facilities. The modelling work in Leeds was supported by the NERC SISLAC grant NE/R001782/1. MPC is also supported by a Royal Society Wolfson Merit award.

We thank Dimitry Belikov, Satoshi Sugawara, Shigeyuki Ishidoya and colleagues for sharing the NIES TM modelling results

495 and stratospheric observations from the balloon sampling-system with us.

13. References

Andrews, A. E., Boering, K. A., Daube, B. C., Wofsy, S. C., Hintsa, E. J., Weinstock, E. M. and Bui, T. P.: Empirical age spectra for the lower tropical stratosphere from in situ observations of CO2: Implications for stratospheric transport, Journal of Geophysical Research Atmospheres, 104(D21), 26581–26595, doi:10.1029/1999JD900150, 1999.

500 Andrews, A. E., Boering, K. A., Wofsy, S. C., Daube, B. C., Jones, D. B., Alex, S., Loewenstein, M., Podolske, J. R. and Strahan, S. E.: Empirical age spectra for the midlatitude lower stratosphere from in situ observations of CO2: Quantitative evidence for a subtropical "barrier" to horizontal transport, Journal of Geophysical Research Atmospheres, 106(D10), 10257–10274, doi:10.1029/2000JD900703, 2001.

Baggenstos, D., Häberli, M., Schmitt, J., Shackleton, S. A., Birner, B., Severinghaus, J. P., Kellerhals, T. and Fischer, H.:

505 Earth's radiative imbalance from the Last Glacial Maximum to the present, Proceedings of the National Academy of Sciences of the United States of America, 116, 14881–14886, doi:10.1073/pnas.1905447116, 2019.

Baldwin, M. P., Gray, L. J., Dunkerton, T. J., Hamilton, K., Haynes, P. H., Randel, W. J., Holton, J. R., Alexander, M. J., Hirota, I., Horinouchi, T., Jones, D. B. A., Kinnersley, J. S., Marquardt, C., Sato, K. and Takahashi, M.: The quasi-biennial oscillation, Reviews of Geophysics, 39, 179–229, doi:10.1029/1999RG000073, 2001.

510 Banks, P. M. and Kockarts, G.: Aeronomy, Part B, 1st ed., Academic Press, New York and London., 1973.

Belikov, D., Sugawara, S., Ishidoya, S., Hasebe, F., Maksyutov, S., Aoki, S., Morimoto, S. and Nakazawa, T.: Three-dimensional simulation of stratospheric gravitational separation using the NIES global atmospheric tracer transport model, Atmospheric Chemistry and Physics, 19, 5349–5361, doi:10.5194/acp-19-5349-2019, 2019.

Belikov, D. A., Maksyutov, S., Sherlock, V., Aoki, S., Deutscher, N. M., Dohe, S., Griffith, D., Kyro, E., Morino, I., Nakazawa,

515 T., Notholt, J., Rettinger, M., Schneider, M., Sussmann, R., Toon, G. C., Wennberg, P. O. and Wunch, D.: Simulations of column-averaged CO2 and CH4 using the NIES TM with a hybrid sigma-isentropic (σ-θ) vertical coordinate, Atmospheric Chemistry and Physics, 13, 1713–1732, doi:10.5194/acp-13-1713-2013, 2013.

Bent, J. D.: Airborne Oxygen Measurements over the Southern Ocean as an Integrated Constraint of Seasonal Biogeochemical Processes, University of California San Diego., 2014.

520 Bereiter, B., Shackleton, S., Baggenstos, D., Kawamura, K. and Severinghaus, J.: Mean global ocean temperatures during the last glacial transition, Nature, 553, 39–44, doi:10.1038/nature25152, 2018.

Bieri, R. H. and Koide, M.: Noble gases in the atmosphere between 43 and 63 kilometers, Journal of Geophysical Research, 75(33), 6731–6735, doi:10.1029/jc075i033p06731, 1970.

Blaine, T. W., Keeling, R. F. and Paplawsky, W. J.: An improved inlet for precisely measuring the atmospheric Ar/N_2 ratio,
525 Atmospheric Chemistry and Physics, 6, 1181–1184, doi:10.5194/acp-6-1181-2006, 2006.

Boering, K. A., Wofsy, S. C., Daube, B. C., Schneider, H. R., Loewenstein, M., Podolske, J. R. and Conway, T. J.: Stratospheric mean ages and transport rates from observations of carbon dioxide and nitrous oxide, Science, 274(5291), 1340–1343, doi:10.1126/science.274.5291.1340, 1996.

Bönisch, H., Engel, A., Curtius, J., Birner, T. and Hoor, P.: Quantifying transport into the lowermost stratosphere using
530 simultaneous in-situ measurements of SF_6 and CO_2, Atmospheric Chemistry and Physics, 9, 5905–5919, doi:10.5194/acp-9-5905-2009, 2009.

Brewer, A. W.: Evidence for a world circulation provided by the measurements of helium and water vapour distribution in the stratosphere, Quarterly Journal of the Royal Meteorological Society, 75(326), 351–363, doi:10.1002/qj.49707532603, 1949.

Butchart, N.: The Brewer-Dobson circulation, Reviews of Geophysics, 52, 157–184, doi:10.1002/2013RG000448, 2014.

535 Chabrillat, S., Vigouroux, C., Christophe, Y., Engel, A., Errera, Q., Minganti, D., Monge-Sanz, B. M., Segers, A. and Mahieu, E.: Comparison of mean age of air in five reanalyses using the BASCOE transport model, Atmospheric Chemistry and Physics, 18, 14715–14735, doi:10.5194/acp-18-14715-2018, 2018.

Chapman, S., Cowling, T. G. and Burnett, D.: The mathematical theory of non-uniform gases: an account of the kinetic theory of viscosity, thermal conduction and diffusion in gases, 3rd ed., Cambridge University Press., 1990.

540 Chipperfield, M. P.: New version of the TOMCAT/SLIMCAT off-line chemical transport model: Intercomparison of stratospheric tracer experiments, Quarterly Journal of the Royal Meteorological Society, 132(617), 1179–1203, doi:10.1256/qj.05.51, 2006.

Chipperfield, M. P., Bekki, S., Dhomse, S., Harris, N. R. P., Hassler, B., Hossaini, R., Steinbrecht, W., Thiéblemont, R. and Weber, M.: Detecting recovery of the stratospheric ozone layer, Nature, 549, 211–218, doi:10.1038/nature23681, 2017.

545 Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M.,

Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., Mcnally, A. P., Monge-Sanz, B. M., Morcrette, J. J., Park, B. K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J. N. and Vitart, F.: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system, Quarterly Journal of

550   the Royal Meteorological Society, 137, 553–597, doi:10.1002/qj.828, 2011.

Diskin, G. S., Podolske, J. R., Sachse, G. W. and Slate, T. A.: Open-path airborne tunable diode laser hygrometer, in Diode Lasers and Applications in Atmospheric Sensing, edited by A. Fried, pp. 196–204, Proceedings of the Society of Photo-Optical Instrumentation Engineers (SPIE)., 2002.

Dobson, G. M. B.: Origin and distribution of the polyatomic molecules in the atmosphere, Proceedings of the Royal Society

555   of London. Series A. Mathematical and Physical Sciences, 236, 187–193, doi:10.1098/rspa.1956.0127, 1956.

Engel, A., Möbius, T., Bönisch, H., Schmidt, U., Heinz, R., Levin, I., Atlas, E., Aoki, S., Nakazawa, T., Sugawara, S., Moore, F., Hurst, D., Elkins, J., Schauffler, S., Andrews, A. and Boering, K.: Age of stratospheric air unchanged within uncertainties over the past 30 years, Nature Geoscience, 2, 28–31, doi:10.1038/ngeo388, 2009.

Engel, A., Bönisch, H., Ullrich, M., Sitals, R., Membrive, O., Danis, F. and Crevoisier, C.: Mean age of stratospheric air

560   derived from AirCore observations, Atmospheric Chemistry and Physics, 17, 6825–6838, doi:10.5194/acp-17-6825-2017, 2017.

Flury, T., Wu, D. L. and Read, W. G.: Variability in the speed of the Brewer-Dobson circulation as observed by Aura/MLS, Atmospheric Chemistry and Physics, 13, 4563–4575, doi:10.5194/acp-13-4563-2013, 2013.

Green, P. J.: Reversible jump Markov chain monte carlo computation and Bayesian model determination, Biometrika, 82(4),

565   711–732, doi:10.1093/biomet/82.4.711, 1995.

Hall, T. M. and Plumb, R. A.: Age as a diagnostic of stratospheric transport, Journal of Geophysical Research, 99(D1), 1059–1070, doi:10.1029/93JD03192, 1994.

Hastings, W. K.: Monte carlo sampling methods using Markov chains and their applications, Biometrika, 57(1), 97–109, doi:10.1093/biomet/57.1.97, 1970.

570   Hauck, M., Fritsch, F., Garny, H. and Engel, A.: Deriving stratospheric age of air spectra using an idealized set of chemically active trace gases, Atmospheric Chemistry and Physics, 19, 5269–5291, doi:10.5194/acp-19-5269-2019, 2019.

Hauck, M., Bönisch, H., Hoor, P., Keber, T., Ploeger, F., Schuck, T. J. and Engel, A.: A convolution of observational and model data to estimate age of air spectra in the northern hemispheric lower stratosphere, Atmospheric Chemsitry and Physics Discussions, in review, doi:10.5194/acp-2020-167, 2020.

575   Haynes,   P.:   Stratospheric   Dynamics,   Annual   Review   of   Fluid   Mechanics,   37,   263–293, doi:10.1146/annurev.fluid.37.061903.175710, 2005.

Headly, M. A. and Severinghaus, J. P.: A method to measure Kr/N2 ratios in air bubbles trapped in ice cores and its application in reconstructing past mean ocean temperature, Journal of Geophysical Research, 112(D19105), 1–12, doi:10.1029/2006JD008317, 2007.

580   Holton, J. R., Haynes, P. H., Mcintyre, M. E., Douglass, A. R. and Rood, B.: Stratosphere-Troposphere, Review of Geophysics, (95), 403–439, 1995.

Hoor, P., Gurk, C., Brunner, D., Hegglin, M. I., Wernli, H. and Fischer, H.: Seasonality and extent of extratropical TST derived from in-situ CO measurements during SPURT, Atmospheric Chemistry and Physics Discussions, 4, 1427–1442, doi:10.5194/acpd-4-1691-2004, 2004.

585   Ishidoya, S., Sugawara, S., Morimoto, S., Aoki, S. and Nakazawa, T.: Gravitational separation of major atmospheric components of nitrogen and oxygen in the stratosphere, Geophysical Research Letters, 35(L03811), doi:10.1029/2007GL030456, 2008.

Ishidoya, S., Sugawara, S., Morimoto, S., Aoki, S., Nakazawa, T., Honda, H. and Murayama, S.: Gravitational separation in the stratosphere - A new indicator of atmospheric circulation, Atmospheric Chemistry and Physics, 13, 8787–8796, 590   doi:10.5194/acp-13-8787-2013, 2013.

Ishidoya, S., Sugawara, S., Inai, Y., Morimoto, S., Honda, H., Ikeda, C., Hashida, G., Machida, T., Tomikawa, Y., Toyoda, S., Goto, D., Aoki, S. and Nakazawa, T.: Gravitational separation of the stratospheric air over Syowa, Antarctica and its connection with meteorological fields, Atmospheric Science Letters, 19, 1–7, doi:10.1002/asl.857, 2018.

Karion, A., Sweeney, C., Wolter, S., Newberger, T., Chen, H., Andrews, A., Kofler, J., Neff, D. and Tans, P.: Long-term 595   greenhouse gas measurements from aircraft, Atmospheric Measurement Techniques, 6, 511–526, doi:10.5194/amt-6-511-2013, 2013.

Keeling, C. D., Stephen, C., Piper, S. C., Bacastow, R. B., Wahlen, M., Whorf, T. P., Heimann, M. and Meijer, H. A.: Exchanges of atmospheric CO2 and 13CO2 with the terrestrial biosphere and oceans from 1978 to 2000, Global Aspects, SIO Reference Series, 01–06, 83–113, 2001.

600   Keeling, R. F.: Development of an interferometric oxygen analyzer for precise measurement of the atmospheric O2 mole fraction, Harvard University., 1988.

Keeling, R. F., Blaine, T., Paplawsky, B., Katz, L., Atwood, C. and Brockwell, T.: Measurement of changes in atmospheric Ar/N_2 ratio using a rapid-switching, single-capillary mass spectrometer system, Tellus, 56B, 322–338, doi:10.1111/j.1600-0889.2004.00117.x, 2004.

605   Kovács, T., Feng, W., Totterdill, A., Plane, J. M. C., Dhomse, S., Gómez-Martín, J. C., Stiller, G. P., Haenel, F. J., Smith, C., Forster, P. M., García, R. R., Marsh, D. R. and Chipperfield, M. P.: Determination of the atmospheric lifetime and global

warming potential of sulfur hexafluoride using a three-dimensional model, Atmospheric Chemistry and Physics, 17, 883–898, doi:10.5194/acp-17-883-2017, 2017.

610 Krol, M., De Bruine, M., Killaars, L., Ouwersloot, H., Pozzer, A., Yin, Y., Chevallier, F., Bousquet, P., Patra, P., Belikov, D., Maksyutov, S., Dhomse, S., Feng, W. and Chipperfield, M. P.: Age of air as a diagnostic for transport timescales in global models, Geoscientific Model Development, 11, 3109–3130, doi:10.5194/gmd-11-3109-2018, 2018.

Li, F., Waugh, D. W., Douglass, A. R., Newman, P. A., Pawson, S., Stolarski, R. S., Strahan, S. E. and Nielsen, J. E.: Seasonal variations of stratospheric age spectra in the Goddard Earth Observing System Chemistry Climate Model (GEOSCCM), Journal of Geophysical Research Atmospheres, 117(D05134), 1–14, doi:10.1029/2011JD016877, 2012.

615 Linz, M., Plumb, R. A., Gerber, E. P. and Sheshadri, A.: The relationship between age of air and the diabatic circulation of the stratosphere, Journal of the Atmospheric Sciences, 73, 4507–4518, doi:10.1175/JAS-D-16-0125.1, 2016.

Linz, M., Plumb, R. A., Gerber, E. P., Haenel, F. J., Stiller, G., Kinnison, D. E., Ming, A. and Neu, J. L.: The strength of the meridional overturning circulation of the stratosphere, Nature Geoscience, 10, 663–667, doi:10.1038/ngeo3013, 2017.

Malinverno, A.: Parsimonious Bayesian Markov chain Monte Carlo inversion in a nonlinear geophysical problem, Geophysics 620 Journal International, 151, 675–688, 2002.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H. and Teller, E.: Equation of state calculations by fast computing machines, The Journal of Chemical Physics, 21(6), 1087–1092, doi:10.1063/1.1699114, 1953.

Monge-Sanz, B. M., Chipperfield, M. P., Simmons, A. J. and Uppala, S. M.: Mean age of air and transport in a CTM: Comparison of different ECMWF analyses, Geophysical Research Letters, 34(L04801), 1–5, doi:10.1029/2006GL028515, 625 2007.

Montzka, S. A., Dutton, G. S., Yu, P., Ray, E., Portmann, R. W., Daniel, J. S., Kuijpers, L., Hall, B. D., Mondeel, D., Siso, C., Nance, J. D., Rigby, M., Manning, A. J., Hu, L., Moore, F., Miller, B. R. and Elkins, J. W.: An unexpected and persistent increase in global emissions of ozone-depleting CFC-11, Nature, 557, 413–417, doi:10.1038/s41586-018-0106-2, 2018.

Morgan, E. J., Stephens, B. B., Birner, B. and Keeling, R. F.: Observational Constraints on the Vertical and Interhemispheric 630 Gradient of the Ar/N2 Ratio in the Troposphere, in Abstract CT14A-0845 presented at Ocean Sciences Meeting 2020, AGU, San Diego., 2020.

Plumb, R. A.: Tracer interrelationships in the stratosphere, Reviews of Geophysics, 45(RG4005), 1–33, doi:10.1029/2005RG000179, 2007.

Podglajen, A. and Ploeger, F.: Retrieving the age of air spectrum from tracers: Principle and method, Atmospheric Chemistry 635 and Physics, 19, 1767–1783, doi:10.5194/acp-19-1767-2019, 2019.

Prather, M. J., Zhu, X., Flynn, C. M., Strode, S. A., Rodriguez, J. M., Steenrod, S. D., Liu, J., Lamarque, J. F., Fiore, A. M., Horowitz, L. W., Mao, J., Murray, L. T., Shindell, D. T. and Wofsy, S. C.: Global atmospheric chemistry - Which air matters, Atmospheric Chemistry and Physics, 17, 9081–9102, doi:10.5194/acp-17-9081-2017, 2017.

Ray, E. A., Moore, F. L., Elkins, J. W., Dutton, G. S., Fahey, D. W., Vömel, H., Oltmans, S. J. and Rosenlof, K. H.: Transport into the Northern Hemisphere lowermost stratosphere revealed by in situ tracer measurements, Journal of Geophysical Research Atmospheres, 104(D21), 26565–26580, doi:10.1029/1999JD900323, 1999.

Ray, E. A., Moore, F. L., Rosenlof, K. H., Davis, S. M., Boenisch, H., Morgenstern, O., Smale, D., Rozanov, E., Hegglin, M., Pitari, G., Mancini, E., Braesicke, P., Butchart, N., Hardiman, S., Li, F., Shibata, K. and Plummer, D. A.: Evidence for changes in stratospheric transport and mixing over the past three decades based on multiple data sets and tropical leaky pipe analysis, Journal of Geophysical Research Atmospheres, 115(D21304), 1–16, doi:10.1029/2010JD014206, 2010.

Ray, E. A., Moore, F. L., Rosenlof, K. H., Davis, S. M., Sweeney, C., Tans, P., Wang, T., Elkins, J. W., Bönisch, H., Engel, A., Sugawara, S., Nakazawa, T. and Aoki, S.: Improving stratospheric transport trend analysis based on SF6 and CO2 measurements, Journal of Geophysical Research: Atmospheres, 119, 14110–14128, doi:10.1002/2014JD021802, 2014.

Reid, R. C., Prausnitz, J. M. and Poling, B. E.: The properties of gases and liquids, 4th ed., edited by B. Sun and G. H. Fleck, McGraw-Hill, New York., 1987.

Ritz, S. P., Stocker, T. F. and Severinghaus, J. P.: Noble gases as proxies of mean ocean temperature: Sensitivity studies using a climate model of reduced complexity, Quaternary Science Reviews, 30, 3728–3741, doi:10.1016/j.quascirev.2011.09.021, 2011.

Salby, M. L. and Callaghan, P. F.: Influence of the Brewer-Dobson circulation on stratosphere-troposphere exchange, Journal of Geophysical Research Atmospheres, 111(21), 1–9, doi:10.1029/2006JD007051, 2006.

Santoni, G. W., Daube, B. C., Kort, E. A., Jiménez, R., Park, S., Pittman, J. V., Gottlieb, E., Xiang, B., Zahniser, M. S., Nelson, D. D., Mcmanus, J. B., Peischl, J., Ryerson, T. B., Holloway, S., Andrews, A. E., Sweeney, C., Hall, B., Hintsa, E. J., Moore, F. L., Elkins, J. W., Hurst, D. F., Stephens, B. B., Bent, J. and Wofsy, S. C.: Evaluation of the airborne quantum cascade laser spectrometer (QCLS) measurements of the carbon and greenhouse gas suite - CO2, CH4, N2O, and CO - during the CalNex and HIPPO Campaigns, Atmospheric Measurement Techniques, 7, 1509–1526, doi:10.5194/amt-7-1509-2014, 2014.

Schoeberl, M. R., Douglass, A. R., Polansky, B., Bonne, C., Walker, K. A. and Bernath, P.: Estimation of stratospheric age spectrum from chemical tracers, Journal of Geophysical Research Atmospheres, 110(D21303), 1–18, doi:10.1029/2005JD006125, 2005.

Seltzer, A. M., Ng, J., Danskin, W. R., Kulongoski, J. T., Gannon, R. S., Stute, M. and Severinghaus, J. P.: Deglacial water-table decline in Southern California recorded by noble gas isotopes, Nature Communications, 10, 1–6, doi:10.1038/s41467-

019-13693-2, 2019.

Škerlak, B., Sprenger, M. and Wernli, H.: A global climatology of stratosphere-troposphere exchange using the ERA-Interim data set from 1979 to 2011, Atmospheric Chemistry and Physics, 14, 913–937, doi:10.5194/acp-14-913-2014, 2014.

Stephens, B. B.: ORCAS Merge Products. Version 1.0, , doi:10.5065/D6SB445X, 2017.

670 Stephens, B. B., Long, M. C., Keeling, R. F., Kort, E. A., Sweeney, C., Apel, E. C., Atlas, E. L., Beaton, S., Bent, J. D., Blake, N. J., Bresch, J. F., Casey, J., Daube, B. C., Diao, M., Diaz, E., Dierssen, H., Donets, V., Gao, B. C., Gierach, M., Green, R., Haag, J., Hayman, M., Hills, A. J., Hoecker-Martínez, M. S., Honomichl, S. B., Hornbrook, R. S., Jensen, J. B., Li, R. R., McCubbin, I., McKain, K., Morgan, E. J., Nolte, S., Powers, J. G., Rainwater, B., Randolph, K., Reeves, M., Schauffler, S. M., Smith, K., Smith, M., Stith, J., Stossmeister, G., Toohey, D. W. and Watt, A. S.: The O2/N2 ratio and CO2 airborne

675 Southern Ocean study, Bulletin of the American Meteorological Society, 99, 381–402, doi:10.1175/BAMS-D-16-0206.1, 2018.

Strahan, S. E., Douglass, A. R., Stolarski, R. S., Akiyoshi, H., Bekki, S., Braesicke, P., Butchart, N., Chipperfield, M. P., Cugnet, D., Dhomse, S., Frith, S. M., Gettelman, A., Hardiman, S. C., Kinnison, D. E., Lamarque, J. F., Mancini, E., Marchand, M., Michou, M., Morgenstern, O., Nakamura, T., Olivié, D., Pawson, S., Pitari, G., Plummer, D. A., Pyle, J. A., Scinocca, J.

680 F., Shepherd, T. G., Shibata, K., Smale, D., Teyssèdre, H., Tian, W. and Yamashita, Y.: Using transport diagnostics to understand chemistry climate model ozone simulations, Journal of Geophysical Research Atmospheres, 116(17), 1–18, doi:10.1029/2010JD015360, 2011.

Sugawara, S., Ishidoya, S., Aoki, S., Morimoto, S., Nakazawa, T., Toyoda, S., Inai, Y., Hasebe, F., Ikeda, C., Honda, H., Goto, D. and Putri, F. A.: Age and gravitational separation of the stratospheric air over Indonesia, Atmospheric Chemistry and

685 Physics, 18, 1819–1833, doi:10.5194/acp-18-1819-2018, 2018.

Taylor, R. and Krishna, R.: Multicomponent mass transfer, 2nd ed., John Wiley & Sons., 1993.

Waldmann, L.: Die Temperaturerscheinungen bei der Diffusion in ruhenden Gasen und ihre meßtechnische Anwendung, Zeitschrift für Physik, 124(2), 2–29, doi:10.1007/BF01374919, 1947.

Waugh, D.: Atmospheric dynamics: The age of stratospheric air, Nature Geoscience, 2, 14–16, doi:10.1038/ngeo397, 2009.

690 Waugh, D. W. and Hall, T. M.: Age of stratospheric air: Theory, observations, and models, Reviews of Geophysics, 40(4), doi:10.1029/2000RG000101, 2002.

Wofsy, S. C., the HIPPO science team, cooporating modelers and satellite teams: HIAPER Pole-to-Pole Observations (HIPPO): Fine-grained, global-scale measurements of climatically important atmospheric gases and aerosols, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 369, 2073–2086,

695 doi:10.1098/rsta.2010.0313, 2011.

Wofsy, S. C., Daube, B. C., Jimenez, R., Kort, E., Pittman, J. V., Park, S., Commane, R., Xiang, B., Santoni, G., Jacob, D., Fisher, J., Pickett-Heaps, C., Wang, H., Wecht, K., Wang, Q.-Q., Stephens, B. B., Shertz, S., Watt, A. S., Romashkin, P., Campos, T., Haggerty, J., Cooper, W. A., Rogers, D., Beaton, S., Hendershot, R., Elkins, J. W., Fahey, D. W., Gao, R. S., Moore, F., Montzka, S. A., Schwarz, J. P., Perring, A. E., Hurst, D., Miller, B. R., Sweeney, C., Oltmans, S., Nance, D., Hintsa, E., Dutton, G., Watts, L. A., Spackman, J. R., Rosenlof, K. H., Ray, E. A., Hall, B., Zondlo, M. A., Diao, M., Keeling, R., Bent, J., Atlas, E. L., Lueb, R. and Mahoney, M. J.: HIPPO MEDUSA Flask Sample Trace Gas and Isotope Data (R_20121129), , doi:10.3334/CDIAC/HIPPO_014, 2012.

Wofsy, S. C., Afshar, S., Allen, H. M., Apel, E., Asher, E. C., Barletta, B., Bent, J., Bian, H., Biggs, B. C., Blake, D. R., Blake, N., Bourgeois, I., Brock, C. A., Brune, W. H., Budney, J. W., Bui, T. P., Butler, A., Campuzano-Jost, P., Chang, C. S., Chin, M., Commane, R., Correa, G., Crounse, J. D., Cullis, P. D., Daube, B. C., Day, D. A., Dean-Day, J. M., Dibb, J. E., Digangi, J. P., Diskin, G. S., Dollner, M., Elkins, J. W., Erdesz, F., Fiore, A. M., Flynn, C. M., Froyd, K., Gesler, D. W., Hall, S. R., Hanisco, T. F., Hannun, R. A., Hills, A. J., Hintsa, E. J., Hoffman, A., Hornbrook, R. S., Huey, L. G., Hughes, S., Jimenez, J. L., Johnson, B. J., Katich, J. M., Keeling, R., Kim, M. J., Kupc, A., Lait, L. R., Lamarque, J. F., Liu, J., McKain, K., McLaughlin, R. J., Meinardi, S., Miller, D. O., Montzka, S. A., Moore, F. L., Morgan, E. J., Murphy, D. M., Murray, L. T., Nault, B. A., Neuman, J. A., Newman, P. A., Nicely, J. M., Pan, X., Paplawsky, W., Peischl, J., Prather, M. J., Price, D. J., Ray, E., Reeves, J. M., Richardson, M., Rollins, A. W., Rosenlof, K. H., Ryerson, T. B., Scheuer, E., Schill, G. P., Schroder, J. C., Schwarz, J. P., St.Clair, J. M., Steenrod, S. D., Stephens, B. B., Strode, S. A., Sweeney, C., Tanner, D., Teng, A. P., Thames, A. B., Thompson, C. R., Ullmann, K., Veres, P. R., Vizenor, N., Wagner, N. L., Watt, A., Weber, R., Weinzierl, B., et al.: ATom: Merged Atmospheric Chemistry, Trace Gases, and Aerosols, , doi:10.3334/ornldaac/1581, 2018.

Zondlo, M. A., Paige, M. E., Massick, S. M. and Silver, J. A.: Vertical cavity laser hygrometer for the National Science Foundation Gulfstream-V aircraft, Journal of Geophysical Research Atmospheres, 115(D20309), 1–14, doi:10.1029/2010JD014445, 2010.
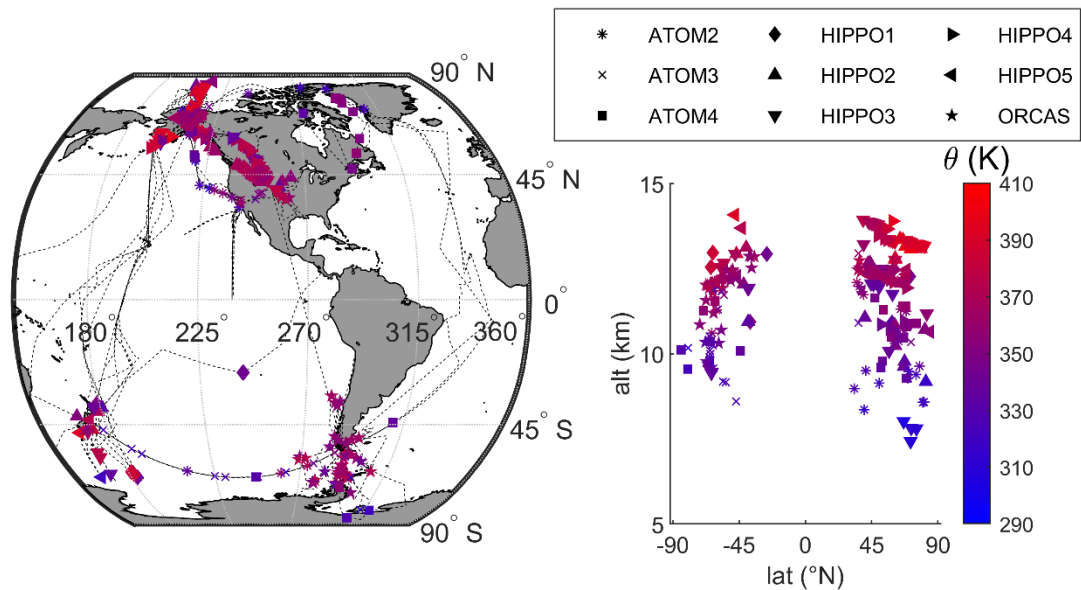
Figure 1. Horizontal (left) and vertical (right) distribution of airborne flask sample locations identified as being of stratospheric origin (see text). Thin dashed black lines on the map illustrate the flight tracks of all 9 campaigns. Symbols indicate the campaigns during which the stratospheric samples were collected, and colours show the potential temperature at which the sample was taken.
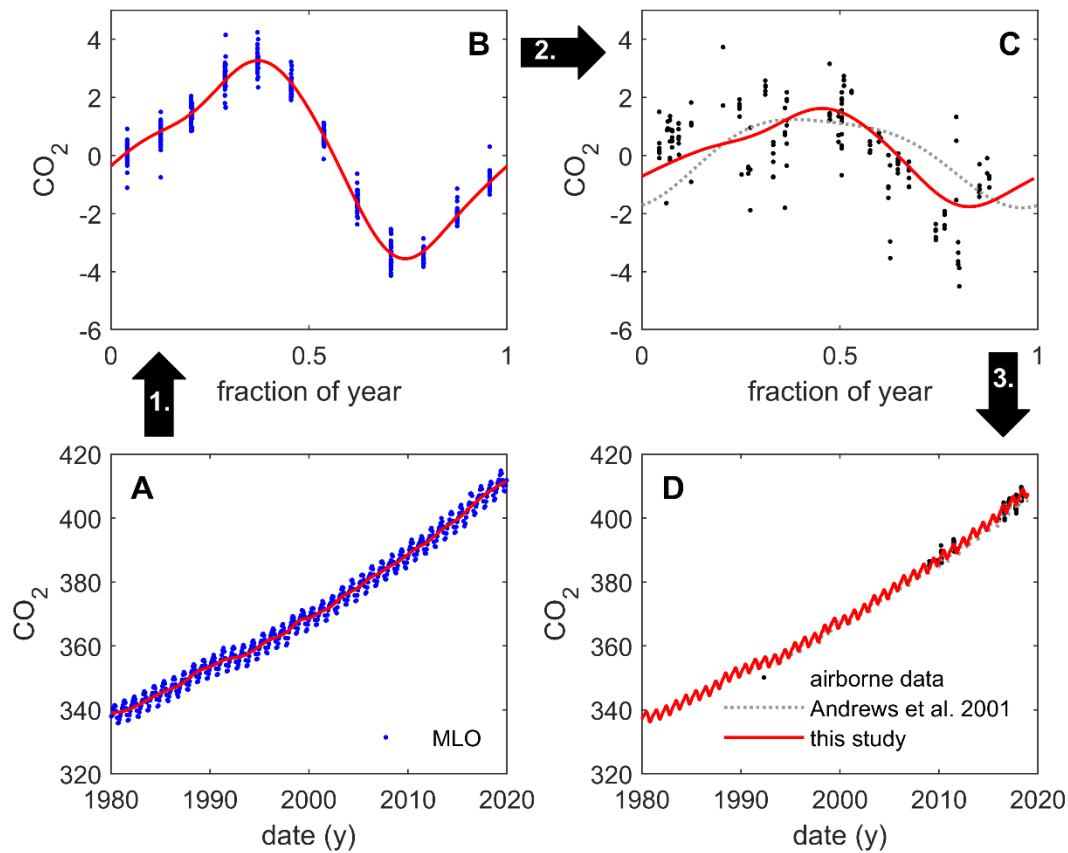
Figure 2. Illustration of the major steps in deriving a tropical upper tropospheric reference time series. Panel (a) shows the monthly mean surface CO_2 concentrations at Mauna Loa (MLO) (Keeling et al., 2001) with a stiff spline trend shown in red. Panel (b) shows a fit of the mean seasonal cycle at MLO compared to detrended observations from 1980–2019. Panel (c) shows the 1-month-lagged seasonal cycle (red line) rescaled in amplitude to match detrended airborne observations in the equatorial upper troposphere ($20°S<$ lat $< 20°N$, alt >8 km, black). The seasonal cycle derived by Andrews et al. (2001) is presented as a dotted grey line for reference. Panel (d) shows the resulting upper tropospheric mean reference time series (TRTS, red) used for CO_2 in the age of air algorithm with the airborne observations from the tropical upper troposphere (black dots) and the time series (grey dotted line) of Andrews et al. (2001).
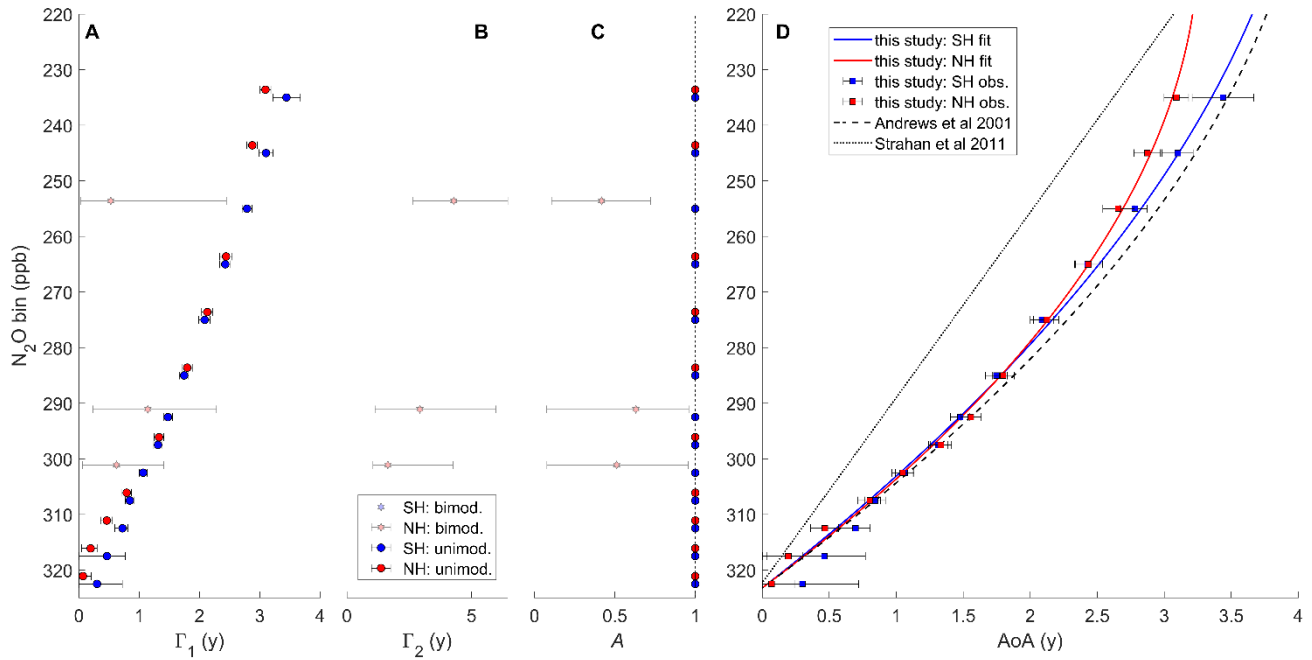
Figure 3. Hemisphere-specific properties of the age spectra in different N$_2$O bins from the 10s-averaged airborne observations estimated by Markov Chain Monte Carlo. Panels (a)-(c) show the value and 95% confidence interval of Γ_1, Γ_2, and A in each bin for either a unimodal (blue and red circles) or bimodal (faint blue and faint red stars) age spectrum depending on which type was preferentially selected by the algorithm. N$_2$O values are offset by 1.4 ppb for all Northern Hemisphere estimates (NH, red) for visual clarity in panel (a)-(c). Panel (d) shows the mean AoA and confidence interval of the age spectra ensemble in each bin (i.e., unimodal and bimodal together). These data are fit by a quadratic polynomial for each hemisphere with a fixed y-intercept (NH AoA = -0.0002381 (323.23 - N$_2$O)^2 + 0.05567 (323.23 - N$_2$O); SH AoA = -0.0001723 (323.23 - N$_2$O)^2 + 0.05317 (323.23 - N$_2$O)). Previous relationships published by Andrews et al. (2001) and Strahan et al. (2011) are given as a dashed and dotted line for reference. Y-intercept values of the previously published N$_2$O-AoA calibrations have been updated to reflect the gradual increase in tropospheric N$_2$O between 1997 and the new reference year 2009.
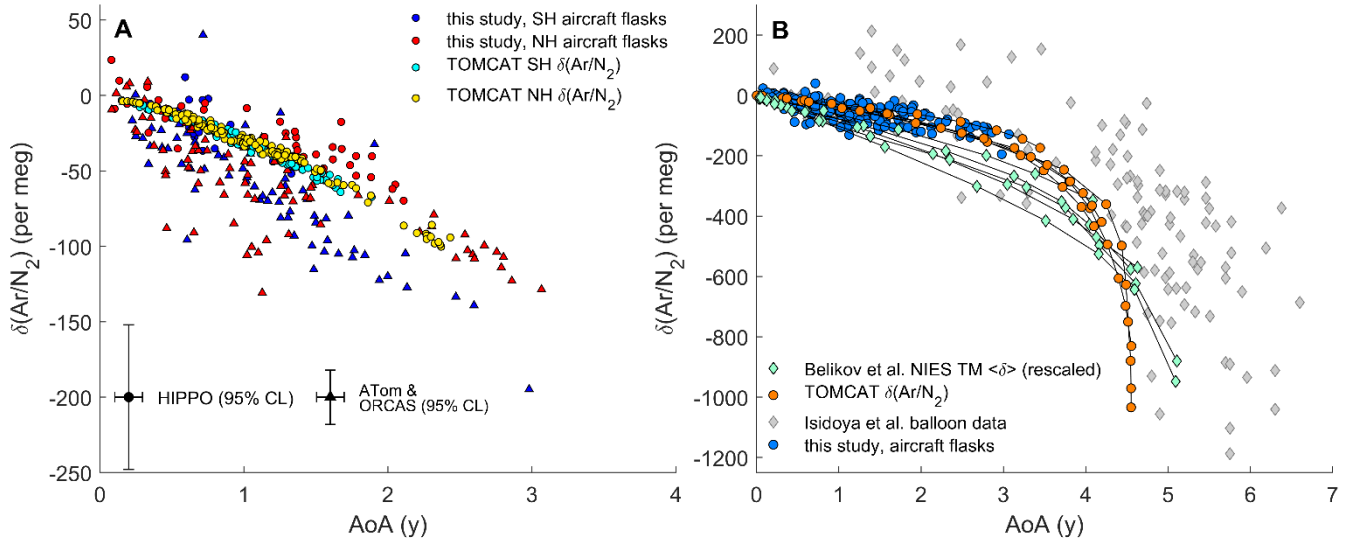
Figure 4. Comparison of age of air (AoA) and gravitational settling (GS) of $\delta(Ar/N_2)$ in models and observations. Panel (a) shows observations from the Northern (red) and Southern (blue) Hemispheres together with TOMCAT model output (SH: cyan; NH: yellow) selected from the closest model grid box in time and space. Because uncertainty for AoA is similar for all stratospheric samples but $\delta(Ar/N_2)$ uncertainty differs between campaigns, representative error bars (95% confidence interval) are shown for HIPPO (triangles) and AToM/ORCAS (circles) at an arbitrary AoA. $\delta(Ar/N_2)$ is normalized to yield a delta value of zero in the equatorial free troposphere. In panel (b), observations from airborne campaigns (this study) and the balloon sampling system (Ishidoya et al., 2008, 2013, 2018; Sugawara et al., 2018) are plotted with the AoA-GS relationship observed in TOMCAT and the NIES TM (Belikov et al., 2019) as lines for each of the latitude bands shown in Fig. 5 and points for different altitude bins between 10 and 35 km. To yield an equivalent estimate of $\delta(Ar/N_2)$, $\delta(^{13}CO_2/^{12}CO_2)$ results from the NIES TM (Belikov et al., 2019) have been rescaled according to Eq. (5) and offset by ~10 per meg to account for the different tropospheric reference region in the definition of δ.

760
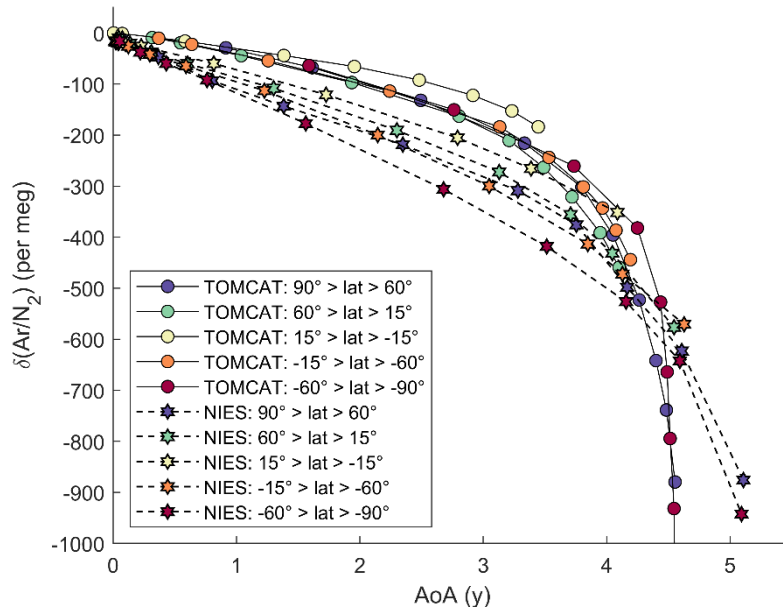
Figure 5. Comparison of the annual and zonal mean age of air (AoA) relationship to gravitational settling simulated in TOMCAT and the NIES TM. The relationship is plotted as lines for the latitude bands indicated by marker colour. For the NIES TM, markers show the vertical profile using all grid boxes available between ~10-35 km. For TOMCAT, markers instead correspond to binned altitude bands between 10 and 35 km with a spacing of 2.5 km because of the finer vertical resolution of the model. To yield an equivalent estimate of $\delta(Ar/N_2)$, $\delta(^{13}CO_2/^{12}CO_2)$ results from the NIES TM (Belikov et al., 2019) have been rescaled according to Eq. (5) and offset by ~10 per meg to account for the different tropospheric reference region in the definition of δ.
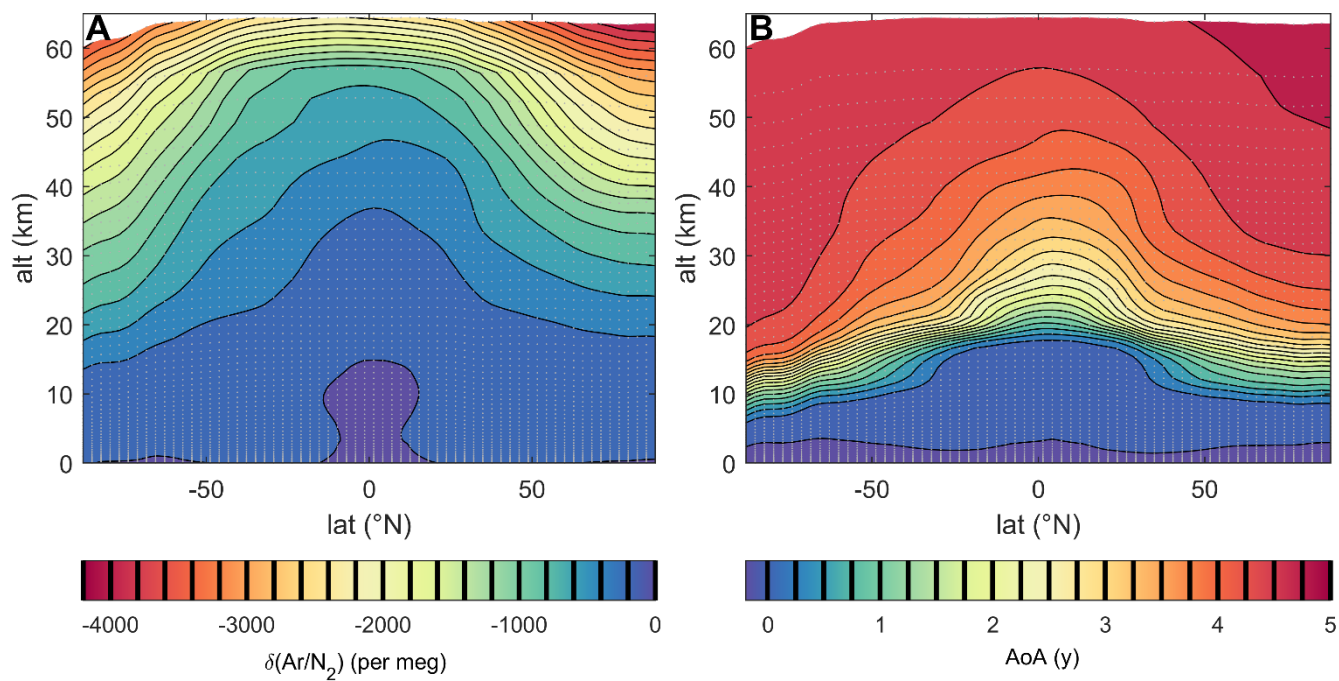
Figure 6. Annual and zonal mean of δ(Ar/N₂) (panel a) and age of air (AoA) (panel b) simulated by TOMCAT. Values of the contour lines are shown as black vertical lines on the colour bar. Grey dots indicate the centre of a grid boxes in TOMCAT. A video of this figure highlighting the natural variability in monthly-mean values is available online with this manuscript.
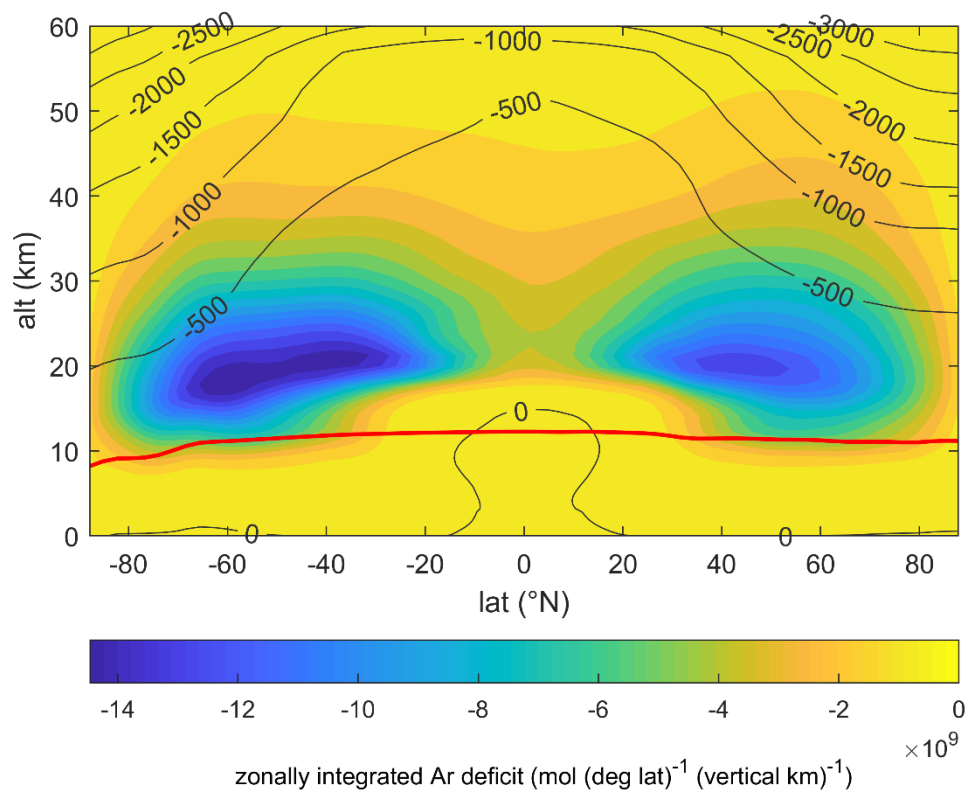
Figure 7. Colormap of annual mean, zonally integrated Ar deficit in TOMCAT. Overlain black contour lines show δ(Ar/N₂) (per meg). The red solid line highlights the mean position of the 200 hPa isobar.

Table B1. Molecular diffusion volumes (Reid et al., 1987) and masses used in this study

Chemical species	Molecular mass (g mol^{-1})	Molecular diffusion volume
Ar	40	16.2
N_2	28	18.5
O_2	32	16.3
air	28.95	19.7

785

Table B2. Characteristic values of variables in Eqs. (B14) and (B15)

Variable/term	Magnitude in extratropics at:			units	Notes or Eq.
	~20 km	~35 km	~57 km		
P	50	5	0.3	mb	
T	215	235	255	K	3
$\dfrac{\partial T}{\partial z}$	3e-04	2.5e-03	2.4e-03	K m^{-1}	3
H_{N2}	6.51e+03	7.11e+03	7.57e+03	m	
H_{Ar}	4.56e+03	4.98e+03	5.30e+03	m	
H_{air}	6.29e+03	6.88e+03	7.32e+03	m	
$\alpha^T_{N2:O2}$		<1.8e-02		-	1
$\alpha^T_{Ar:air}$	~5.6e-02	~6.3e-02	~6.6e-02	-	2
C_{N2}		78.09		%	
C_{O2}		20.95		%	
C_{Ar}		0.93		%	
$\dfrac{\partial C_{N2}}{\partial z}$	2.0e-09	2.6e-09	8.7e-09	m^{-1}	3
$\dfrac{\partial C_{Ar}}{\partial z}$	2.7e-10	3.4e-10	11.6e-10	m^{-1}	3
D^{air}_{N2}	2.3e-04	2.7e-03	5.3e-02	m^{2}s^{-1}	
$D^{air}_{Ar\times(N2,O2)}$	9.8e-09	1.1e-07	2.2e-06	m^{2}s^{-1}	
D^{air}_{Ar}	2.2e-04	2.6e-03	5.0e-02	m^{2}s^{-1}	
$D^{air}_{N2}\dfrac{\partial C_{N2}}{\partial z}$	4.7e-13	7.1e-12	4.6e-10	m s^{-1}	A19
$D^{air}_{N2}\left(\dfrac{1}{H_{N2}}-\dfrac{1}{H_{air}}\right)C_{N2}$	9.6e-10	11.0e-08	1.9e-7	m s^{-1}	A19
$D^{air}_{N2}\alpha^T_{N2:O2}\dfrac{C_{N2}(1-C_{N2})}{T}\dfrac{\partial T}{\partial z}$	<5.9e-12	<8.7e-11	<1.6e-09	m s^{-1}	A19
$D^{air}_{Ar\times(N2,O2)}\dfrac{\partial C_{N2}}{\partial z}$	2.0e-17	3.0e-16	1.9e-14	m s^{-1}	A20
$D^{air}_{Ar\times(N2,O2)}\left(\dfrac{1}{H_{N2}}-\dfrac{1}{H_{air}}\right)C_{N2}$	4.0e-14	4.3e-13	7.8e-12	m s^{-1}	A20
$D^{air}_{Ar\times(N2,O2)}\alpha^T_{N2:O2}\dfrac{C_{N2}(1-C_{N2})}{T}\dfrac{\partial T}{\partial z}$	<2.5e-16	<3.6e-15	<6.8e-14	m s^{-1}	A20
$D^{air}_{Ar}\dfrac{\partial C_{Ar}}{\partial z}$	6.0e-14	8.9e-13	5.8e-11	m s^{-1}	A20
$D^{air}_{Ar}\left(\dfrac{1}{H_{Ar}}-\dfrac{1}{H_{air}}\right)C_{Ar}$	1.3e-10	1.3e-09	2.4e-08	m s^{-1}	A20
$D^{air}_{Ar}\alpha^T_{Ar:air}\dfrac{C_{Ar}}{T}\dfrac{\partial T}{\partial z}$	~9.8-13	~1.6e-11	~3.2e-10	m s^{-1}	A20

[1] 1.8e-02 observed at 293 K (Waldmann, 1947)

[2] assuming $C_{Ar} \approx \dfrac{N_{Ar}}{N_{N2}}$

[3] estimated from TOMCAT results