Numerical Impacts on Tracer Transport: Diagnosing the Influence of Dynamical Core Formulation and Resolution on Stratospheric Transport

AMAN GUPTA,[a,b] EDWIN P. GERBER,[a] R. ALAN PLUMB,[c] AND PETER H. LAURITZEN[d]

[a] *Center for Atmosphere-Ocean Science, Courant Institute of Mathematical Sciences, New York, New York*
[b] *Meteorological Institute Munich, Ludwig-Maximilian University, Munich, Germany*
[c] *Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts*
[d] *National Center for Atmospheric Research, Boulder, Colorado*

(Manuscript received 19 March 2021, in final form 8 July 2021)

ABSTRACT: Accurate representation of stratospheric trace gas transport is important for ozone modeling and climate projection. Intermodel spread can arise from differences in the representation of transport by the diabatic (overturning) circulation versus comparatively faster adiabatic mixing by breaking waves, or through numerical errors, primarily diffusion. This study investigates the impact of these processes on transport using an idealized tracer, the age of air. Transport is assessed in two state-of-the-art dynamical cores based on fundamentally different numerical formulations: finite-volume and spectral element. Integrating the models in free-running and nudged tropical wind configurations reveals the crucial impact of tropical dynamics on stratospheric transport. Using age-budget theory, vertical and horizontal gradients of age allow comparison of the roles of the diabatic circulation, adiabatic mixing, and the numerical diffusive flux. Their respective contribution is quantified by connecting the full 3D model to the tropical leaky pipe framework of Neu and Plumb. Transport by the two cores varies significantly in the free-running integrations, with the age in the middle stratosphere differing by about 2 years primarily due to differences in adiabatic mixing. When winds in the tropics are constrained, the difference in age drops to about 0.5 years; in this configuration, more than half the difference is due to the representation of the diabatic circulation. Numerical diffusion is very sensitive to the resolution of the core, but does not play a significant role in differences between the cores when they are run at comparable resolution. It is concluded that fundamental differences rooted in dynamical core formulation can account for a substantial fraction of transport bias between climate models.

KEYWORDS: Stratospheric circulation; Stratosphere-troposphere coupling; Model comparison; Idealized models; Global transport modeling

1. Introduction

Trace gases and aerosols play a critical role in air quality (Friedl 2007), the radiative balance of our planet, and, in the case of ozone, harmful UV radiation (Molina and Rowland 1974; Eyring et al. 2007). For species with lifetimes shorter than characteristic mixing time scales of the atmosphere, on the order of months in the troposphere to years in the stratosphere, an accurate simulation requires both the representation of sinks and source (emissions and atmospheric chemistry; e.g., Collins et al. 2017) and transport (the advection and mixing of trace gases by the atmospheric circulation; e.g., Plumb 2002). This study focuses on this second factor, building on recent work by Gupta et al. (2020, hereafter referred to as G20) to establish an intercomparison test to assess trace gas transport by the primitive equation solvers, or dynamical cores, of atmospheric models.

G20 explored the impact of model numerics and resolution on transport by the stratosphere–troposphere system, with particular focus on the middle atmosphere. The slow-transport time scales of the Brewer–Dobson circulation (BDC; Butchart 2014), from months in the upper troposphere and lower stratosphere (UTLS) to years in the so-called deep branch of the BDC (Birner and Bönisch 2011), expose the impact of small errors in numerical transport. Trace gas transport by the BDC depends on both a slow overturning of mass from the tropics to the mid- and high latitudes, and a fast, quasi-horizontal mixing of tracers along isentropic (constant potential temperature) surfaces. We refer to the slow overturning as the diabatic circulation, since air must be radiatively warmed in the tropics and cooled in the higher latitudes to ascend or descend through the highly stratified stratosphere, and the faster mixing processes as isentropic mixing, since it is associated with adiabatic transport and mixing. Both components are primarily driven by planetary wave breaking. The westward torque induced by breaking planetary waves forces the mass in the tropics to cross angular momentum surfaces and enter the higher latitudes, the ''Rossby wave pump'' of Holton et al. (1995). Stratospheric transport by a dynamical core depends on its ability to accurately capture both the overturning circulation and isentropic mixing.

G20 extended the dynamical core setup of Held and Suarez (1994) to account for a more active stratosphere (as in Polvani and Kushner 2002) and introduced an idealized tracer, proposing two benchmark tests to assess stratosphere–troposphere coupling and stratospheric transport. The idealized setup and tests, which are reviewed in section 2, strip away all parameterizations of

Supplemental information related to this paper is available at the Journals Online website: https://doi.org/10.1175/JAS-D-21-0085.s1.

Corresponding author: Aman Gupta, ag4680@nyu.edu

DOI: 10.1175/JAS-D-21-0085.1

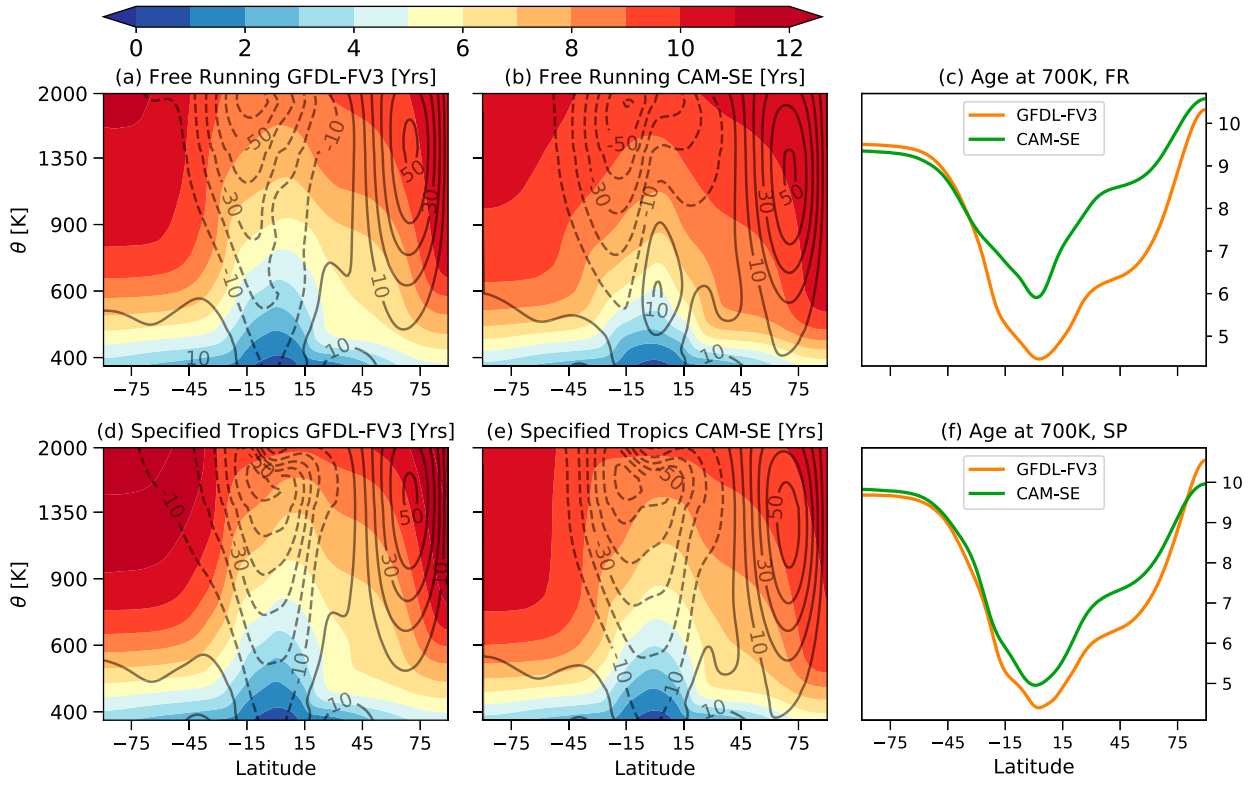

FIG. 1. Zonal-mean age (color; in years) and zonal-mean zonal winds (black contours; in $m\,s^{-1}$) for the (a) free-running GFDL-FV3 "core," (b) free-running CAM-SE core, (d) specified tropics GFDL-FV3 core, and (e) specified tropics CAM-SE "core.". (c),(f) The zonal-mean age at 700-K isentropic level for the FV3 (orange) and CAM-SE (green) core for the (c) free-running and (f) specified tropics integrations. All integrations were run with 1° horizontal resolution and 80 vertical levels.

subgrid-scale processes (including transport and chemistry) to isolate the role of numerics on model climatology, transport, and stratosphere–troposphere coupling. They found differences in transport between four different dynamical cores, including two state-of-the art cores. They concluded that the choice of numerics can significantly impact transport in the stratosphere.

Trace gas transport in G20 was quantified by the age of air (Hall and Plumb 1994; Waugh and Hall 2002), a measure of the transport time scales in the stratosphere discussed in more detail in section 3. The models show large quantitative disagreement in the steady-state age of air profile, which was found to vary by as much as 40% in the winter midlatitude stratosphere in the "free-running" benchmark test, as shown here in Fig. 1. These extreme differences in transport were, however, mostly rooted in differences in the tropical climatology, as opposed to the accuracy of the numerical schemes.

The subtle momentum balance in the tropical stratosphere permits the development of the quasi-biennial oscillation (QBO), a 28-month swing between easterly and westerly jets (Baldwin et al. 2001). Representation of the QBO in comprehensive models depends on both high vertical resolution and the parameterization of unresolved gravity waves (Butchart et al. 2018). The idealized setup of G20 lacks a parameterization of gravity waves, and, as a result, the dynamical cores produced steady climatological wind profiles, which differed greatly between models with spectral- and finite-volume-based numerics. G20 partially resolved this divergence by proposing a second benchmark test in which the tropical wind climatology is constrained (see section 2), but differences between models with finite-volume and spectral numerics persist.

In this study, we assess the intermodel differences in stratospheric transport found in the two G20 benchmark tests. We develop a framework to quantify the fractional contribution in transport differences to the differences in diabatic circulation, isentropic mixing, and numerical diabatic fluxes by applying the theory of age established by Linz et al. (2016, 2021) to estimate the strength of the diabatic circulation and the mixing fluxes in each model. The key is to study the flux of age in isentropic coordinates. Partitioning an isentrope into regions of diabatic upwelling and downwelling allows a convenient separation of the diabatic and adiabatic components of transport, as discussed in section 3.

In section 4, we first estimate the fractional contribution of the diabatic diffusive fluxes to the overall tracer flux, highlighting the importance of high vertical resolution when studying stratospheric transport. In section 5, we establish a framework to quantify the importance of different transport processes on the distribution of mean age using the tropical leaky pipe (TLP) transport model of Neu and Plumb (1999). The framework identifies the role of isentropic mixing

generated in response to the different tropical wind profiles in the free-running experiments. Once the tropical winds (and mixing) are constrained, however, differences in the diabatic circulation between models are most responsible for maintaining differences in large-scale stratospheric transport. Section 6 summarizes our findings.

2. Model integrations and the benchmark tests

G20 proposed two benchmark tests to assess climatological transport in atmospheric general circulation models: a free-running (FR) test and the specified tropical wind (SP) test. The tests were applied to four different dynamical cores with different underlying numerics, developed at two modeling centers, the Geophysical Fluid Dynamics Laboratory (GFDL) and the National Center for Atmospheric Research (NCAR). Here we focus only on two of the four models, the two most recently developed cores, as they capture the key differences observed across all cores. We consider the exact same experiments as in G20, but review the key details here.

a. Model details

The first dynamical core is a finite-volume model based on a cubed sphere grid. It was developed at GFDL and we refer to it as GFDL-FV3, or FV3 for short. The core employs finite-volume schemes in both the vertical and the horizontal to solve the primitive equations. FV3 was built as the core of GFDL's Atmospheric Model, version 3 (AM3; Donner et al. 2011), and a related nonhydrostatic version was recently adopted as the core of the National Centers for Environmental Prediction Global Forecasting System.

The second model is a spectral element core developed at NCAR as part of the Community Atmosphere Model (CAM) framework. It is also based on a cubed sphere grid, but uses a spectral finite element method to discretize the primitive equations in the horizontal, and is coupled to finite-volume Lagrangian advection in the vertical (Lauritzen et al. 2018). We refer to it as CAM-SE, or SE for short. It was designed to become the dynamical core of NCAR's Community Earth System Model, version 2 (CESM2), and used for many of the CESM2 integrations in the Coupled Model Intercomparison Project, phase 6 (CMIP6). Further details on both CAM-SE and GFDL-FV3 are provided in G20 (their section 3 and appendix A).

G20 considered two additional cores, a pseudospectral model developed by GFDL and a finite-volume core based on a latitude–longitude grid, used in the CAM framework. They found, however, that the two finite-volume cores and the two "spectral" cores (i.e., pseudospectral and spectral element) behaved very similarly to each other in terms of their climatological circulation, transport, and sensitivity to resolution.

b. Experiment details

The cores are forced with identical thermal forcings (Held and Suarez 1994; Polvani and Kushner 2002), detailed in section 4a of G20. Newtonian relaxation to an analytic temperature state (which can be interpreted as a state of radiative-convective equilibrium) generates a perpetual northern-winter climatology. The SP test differs from the FR test only in that a Rayleigh damping is applied to the zonal winds of the equatorial stratosphere. As detailed in G20 (their section 4b and appendix B), the damping relaxes the winds above 200 hPa and within 15° of the equator to a steady easterly profile with a time scale of 40 days in the lower stratosphere, rising to 10 days above 3 hPa. The wind profile was chosen to match the winds observed in the finite-volume integrations, which were relatively insensitive to changes in resolution. It should be noted, however, that this choice is not meant to imply that easterlies are the "correct" profile.

To quantify transport, a linearly increasing (in time) clock tracer is introduced near the surface ($p \geq 700$ hPa) as detailed in G20, section 4a. The clock tracer is used to compute the age of air, which provides a measure of transport time scales of trace gases in the stratosphere. The age of air itself can be interpreted as an idealized tracer Γ that "ages" linearly with time, $d\Gamma/dt = 1$, but is reset to zero when it comes in contact with the surface.

As in G20, the models were integrated for 10 000 days and the last 3300 days were used for analysis. To test robustness of transport to model grid resolutions, each core was integrated with at least four different resolutions, as tabulated in Table 1, allowing us to assess the impact of doubling both the horizontal and vertical resolution. For CAM-SE, additional higher resolution runs were considered. For context, the models were integrated with horizontal resolution comparable to that used for comprehensive models in CMIP experiments (approximately 1°), but with vertical resolution that is generally higher than that used for climate integrations, even in the 40 level configurations.

Since the results are less sensitive to the horizontal resolution, we focus on the higher horizontal resolution runs (i.e., 1° × 1°) with 40 and 80 vertical levels, unless otherwise specified.

c. Key results from the G20 benchmark tests

The dynamical cores fail the Held and Suarez (1994) test in the free-running configuration: numerics and resolution fundamentally affect the climatological circulation. As the vertical resolution is increased from 40 to 80 levels, the spectral-based models and the finite-volume-based models develop differences in the climatological state of the tropical stratosphere. With 40 vertical levels, all the models generate tropical easterlies (irrespective of the horizontal resolution) but with 80 levels, the two spectral models generate tropical westerlies (up to 20 m s^{-1}) while the two finite-volume models maintain tropical easterlies (Figs. 1a,b and 2 of G20).

The impact of the tropical wind difference on the overall age profile is significant. The age of air in the free-running model integrations with tropical westerlies (up to 20 m s^{-1}) is up to 40% higher (8.5 years vs 6 years in the winter middle stratosphere) than the models with tropical easterlies (Fig. 1c). The difference in the age of air indicates significant differences in transport. G20 suggested that tropical westerlies induce enhanced tropical–extratropical wave mixing, as the westerlies provide a duct for extratropical Rossby waves to break deep into the tropical atmosphere. The enhanced mixing results in

TABLE 1. Specifications for the dynamical core integrations considered in this study. The models were integrated for 30 years at each resolution for both the free-running (FR) and specified tropics (SP) tests, as described in section 2. The model grids differ slightly from the effective resolution due to the nature of the numerical methods.

Model and numerics	Resolution ID	Grid size	Effective resolution
GFDL-FV3: Cubed-sphere grid, finite-volume formulation	C48L40	192 × 96 × 40	2° × 2°
	C48L80	192 × 96 × 80	
	C90L40	360 × 180 × 40	1° × 1°
	C90L80	360 × 180 × 80	
CAM-SE: Cubed-sphere grid, spectral element formulation	NE16L40	256 × 129 × 40	2° × 2°
	NE16L80	256 × 129 × 80	
	NE30L40	512 × 257 × 40	1° × 1°
	NE30L80	512 × 257 × 80	
	NE30L160	512 × 257 × 160	
	NE60L80	1024 × 513 × 80	0.5° × 0.5°

an increase in the older age transported from the midlatitudes into the tropics. Recirculation of the air results in an older age throughout the stratosphere. This hypothesis was explored with the SP test.

The SP test differs only in one key aspect: the tropical winds are constrained to a specified analytical easterly profile. As a result of this nudging, a closer agreement in tropical winds—and in turn in age of air—was achieved, as seen in Figs. 1d–f: the age difference between CAM-SE and GFDL-FV3 is reduced from 2 years to just 0.5 years in the winter middle stratosphere. The age among the two state-of-the-art cores, however, still exhibits fundamental differences. As detailed in G20 (section 7), the climatological circulation and transport behavior of both cores appears to be stable to further increase in the horizontal and vertical resolution: in a rough sense, they converge to a well-defined state under the SP test. This state, however, differs between the two models and the converged age solutions vary by as much as 15% (~0.5 years) in the winter midlatitudes. The goal of this study is to explore and understand this difference.

Compared to observational estimates (Waugh and Hall 2002), the overall structure of the mean age in our idealized setup is broadly consistent with the mean age of the atmosphere, but biased toward older values. For example, the age in the upper stratosphere exceeds 9 years in the models, compared to 6 years in observations. As explained in G20, the old bias is due to slower transport through the troposphere (there is no convective transport), a lower-than-observed tropopause, and a relatively weak stratospheric diabatic circulation in our idealized model setup.

3. Stratospheric age budget analysis in isentropic coordinates

A schematic of the large-scale stratospheric circulation and transport is shown in Fig. 2. Following Neu and Plumb (1999), we partition the stratosphere into an upwelling (diabatic warming) region in the tropics and downwelling (diabatic cooling) regions in the extratropics. The net overturning, or diabatic circulation, is completed by a flux of mass from the tropics to extratropics, illustrated with the streamlines in the left panel. There is also substantial mixing of air between the upwelling and downwelling regions, illustrated with curly arrows. While this has no impact on the net overturning of mass, it is associated with an adiabatic transport of tracers along isentropes. In the tropical leaky pipe model (Neu and Plumb 1999), the geometry is simplified by lumping the two downwelling regions together (right panel of Fig. 2). The net overturning flow in the meridional, μ_{net}, is viewed as difference between the total entrainment of mass from the tropics to the extratropics, μ_{ent}, and the total mixing flux of air that recirculates back into the tropical pipe, μ_{mix}. The model also allows for the vertical diffusion of age, as represented by vertical, bidirectional curly arrows in the right panel. The goal of this section is to review how the TLP framework can be applied to the three-dimensional circulation of the atmosphere in diabatic coordinates using the age of air, as established by Linz et al. (2016, 2021).

a. Meridional age difference and vertical diffusion

The age of air quantifies the total time elapsed since an air mass was last in contact with the surface (or tropopause, depending on the definition). It represents the mean of a distribution of transit times for air parcels in the air mass. In the stratosphere, the age can be related to the residence times of ozone and ozone destroying substances, most of which have near-surface emissions. It has been extensively used by both observation- and model-based studies as a diagnostic for atmospheric transport (Kida 1983; Waugh et al. 1997; Hall et al. 1999; Engel et al. 2009; Orbe et al. 2013; Linz et al. 2017). Linz et al. (2016) reformulate the TLP in isentropic coordinates to develop a theory to estimate the strength of the diabatic circulation solely from the age of air. Working in potential temperature (or equivalently, entropy) coordinates in the vertical is advantageous primarily because it helps decouple two key transport processes, the slow diabatic advection across isentropes versus fast adiabatic mixing along isentropes (McIntyre and Palmer 1983; Sparling et al. 1997; Haynes and Shuckburgh 2000).

As detailed in section 2 of Linz et al. (2016), for a steady-state circulation, if one neglects vertical diffusion and diabatic dispersion, the age difference between the midlatitude and the tropics is solely determined by the diabatic circulation

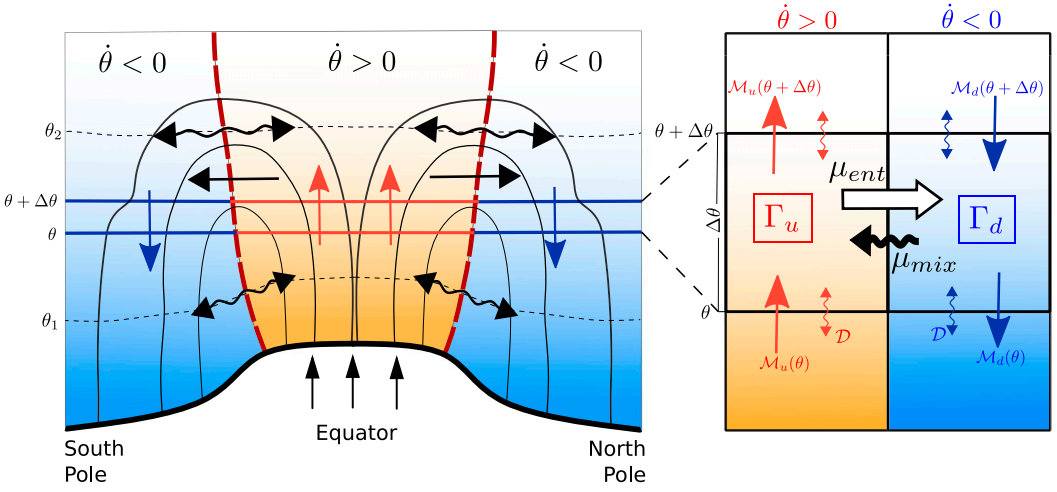

FIG. 2. (left) Zonally averaged and (right) horizontally averaged schematic for stratospheric transport. (left) Red arrows show the diabatic upwelling of mass through the tropical stratosphere in the tropical pipe (orange region), while blue arrows show the diabatic downwelling of mass in the extratropics (blue region). Black arrows show the net flux of mass out of the pipe, while curly arrows illustrate quasi-horizontal mixing of mass along isentropes (dashed black curves). The thin solid black curves show the Lagrangian mean trajectories followed by the air parcels undergoing both slow diabatic advection and rapid isentropic mixing. (right) A simplified schematic of the left panel. The red (blue) arrows denote the upward (downward) diabatic fluxes across the θ isentrope and $(\theta + \Delta\theta)$ isentrope. The vertical curly arrows show diffusive fluxes due to numerical and resolved diffusion. The poleward transport of mass and tracers is represented using the entrainment flux $\mu_{\rm ent}$ (white arrows) and the equatorward transport of tracers is represented using the mixing flux $\mu_{\rm mix}$ (curly black arrow).

strength. The key is to properly define the gross age difference, $\Delta\Gamma = \Gamma_d - \Gamma_u$, where Γ_u and Γ_d are *mass-flux-weighted* representative ages over the upwelling and downwelling partitions of an isentrope, respectively,

$$\Gamma_u(\theta) = \frac{\int_u \rho_\theta \dot\theta \Gamma \, dA}{\int_u \rho_\theta \dot\theta \, dA}, \quad \Gamma_d(\theta) = \frac{\int_d \rho_\theta \dot\theta \Gamma \, dA}{\int_d \rho_\theta \dot\theta \, dA}. \tag{1}$$

Here Γ is the model age, $\dot\theta$ is the diabatic velocity, $\rho_\theta = (1/g)|dp/d\theta|$ is the isentropic density, dA is the infinitesimal area element and $\int_u$ ($\int_d$) represents spatial integration only over the upwelling (downwelling) partition of the θ isentrope, i.e., where $\dot\theta$ is positive (negative). The gross meridional age difference can be directly related to the globally integrated mass above the θ isentrope, $M(\theta)$, and the mass throughput across the isentrope per unit time, $\mathcal{M}(\theta)$:

$$\Delta\Gamma(\theta) = \Gamma_d(\theta) - \Gamma_u(\theta) = \frac{M(\theta)}{\mathcal{M}(\theta)}. \tag{2}$$

The diabatic mass flux $\mathcal{M}$ is the total upward mass flux $\mathcal{M}_u$, or equivalently, the total downward mass flux $\mathcal{M}_d$ assuming steady-state conditions. It is obtained by integrating over the upwelling (downwelling) partition of the stratosphere: $\mathcal{M}_u = \int_u \rho_\theta \dot\theta \, dA$.

The quantity $M(\theta)/\mathcal{M}(\theta)$ (which has units of time) represents the mean residence time (MRT) of air in the atmosphere above the θ isentrope. It presents as an upper bound on the meridional age difference $\Delta\Gamma$ in the limiting case where the age is only transported by diabatic advection and isentropic mixing. Diffusive fluxes, otherwise, act toward reducing $\Delta\Gamma$. In this respect, Eq. (2) can be used to estimate the contribution of unresolved diabatic fluxes and numerical diffusive fluxes to the total age difference $\Delta\Gamma$, as done in section 4. While the theory here is based on a steady state, as detailed in Linz et al. (2016), we account for time variation by computing all quantities from daily data and taking a long average (9 years) to ensure that time dependent terms are negligible.

b. *Vertical age gradient and the tropics–midlatitudes mixing flux*

While the meridional age difference is directly connected to the diabatic circulation strength (Linz et al. 2016), the vertical gradients of the net upwelling age can be used to quantify the efficiency with which planetary wave fluxes transport midlatitude air into the tropics (Linz et al. 2021). In the mid- to upper winter stratosphere, isentropic mixing is effected by the breaking of upward-propagating planetary waves. In the lower stratosphere, this mixing is primarily driven by baroclinic eddies and synoptic-scale wave breaking in both hemispheres (Plumb 2002).

Linz et al. (2021) show that, for a steady circulation in the limit of no vertical diffusion, the vertical age gradient in the tropical pipe, $d\Gamma_u/d\theta$, is the sum of aging by vertical advection and aging by mixing. Mathematically, this is expressed as

$$\frac{d\Gamma_u}{d\theta} = \frac{\sigma_u}{\mathcal{M}} + \mu_{\rm mix}\frac{\Delta\Gamma}{\mathcal{M}}, \tag{3}$$

where σ_u is the isentropic density horizontally integrated over the upwelling partition and $\mu_{\rm mix}$ is the mass flux per kelvin (unit

θ) that mixes the midlatitude air with the tropical air. The first term on the right-hand side captures the net aging of air as it is advected by the diabatic circulation up the tropical pipe: the mass of air per unit potential temperature (kg K^{-1}) divided by the flux of mass (kg s^{-1}): literally, how long it takes the air to rise one unit of potential temperature. The second term captures the total aging of the tropical pipe by older air mixed in from the extratropics: $\Delta\Gamma$ quantifies how much older this extratropical air is, and $\mu_{\rm mix}/\mathcal{M}$ quantifies the relative contribution of this air per unit K. In the special case of vanishing $\Delta\Gamma$ or vanishing $\mu_{\rm mix}$, the air ages solely due to the slow vertical advection of air (Plumb 1996).

Having an a priori knowledge of the age structure and diabatic mass fluxes (i.e., the results from our model integrations), we compute the mixing flux $\mu_{\rm mix}$ as a residual from Eq. (3). It is often transformed to a "mixing efficiency" ε, defined by Neu and Plumb (1999) as the ratio of equatorward mass transport to *net* poleward mass transport. Here, $\mu_{\rm net}$ is the net poleward mass flux, represented by the black poleward arrows in Fig. 2 (left) and is calculated directly from model output.

c. *Applying the theory: An illustration with the GFDL-FV3 core*

To provide a concrete example of the theory, we illustrate the transport metrics from Eqs. (2) and (3) for the 1° L80 free-running GFDL-FV3 core integration in Fig. 3. The metrics will be compared for the two models in the following sections; here we focus on how they are computed and their overall climatological structure. To compute transport metrics in isentropic coordinates using model output on pressure levels, we employ a mass-preserving binning scheme (Yamada and Pauluis 2015) to interpolate from pressure to isentropic coordinates. The details of the scheme are provided as supplementary material.

The daily averaged temperature tendency was used to identify regions of diabatic upwelling and downwelling and the gross ages Γ_u, Γ_d, the mass flux $\mathcal{M}$, the mass M, and the isentropic density ρ_θ were computed daily. All quantities were then averaged over the final 3300 days of the integration (about 9 years). Our choice of a statistically steady climatology without an annual cycle ensured that time averaging first or at the end did not bring any notable difference (not shown). While the theory assumes a steady state, with sufficient time averaging the additional terms that capture temporal variability become negligible. Linz et al. (2016, section 3d, 2021) found that 5 years of averaging was sufficient for high accuracy, even for more realistic simulations with a seasonal cycle.

Figure 3a shows the time-mean zonal-mean age simulated by the finite-volume core. The solid black lines demarcate the mean regions of diabatic upwelling (the tropical region with diabatic velocity $\dot{\theta} > 0$) and diabatic downwelling (higher latitudes with $\dot{\theta} < 0$). Air is youngest at the tropical tropopause, the primary entry point for air into the stratosphere, and older in the extratropics, where mass, on average, descends. The age always increases with potential temperature (i.e., the stratosphere is stratified with respect to age), however, because of mixing.

In the tropics, this stratification reflects the vertical transport of air upward into the stratosphere; the air ages as it is slowly advected up into the tropical atmosphere by the mean overturning circulation. In the extratropics, stratification is maintained by isentropic mixing, which exchanges younger air from the ascending branch of the circulation in the tropics with older descending air in the extratropics. It is for this reason that the weakest vertical gradient is observed in the polar vortex (Northern Hemisphere above 900 K), as strong gradients in potential vorticity associated with the vortex inhibit mixing. The perpetual boreal winter forcing leads to an artificially weak circulation in the austral hemisphere, and the tropical pipe extends all the way to the pole above 1200 K.

We compute the mass-flux-weighted average age over regions of upwelling and downwelling to obtain Γ_u and Γ_d, respectively, plotted in Fig. 3b. The gross ages increase with height, and Γ_d is consistently older than Γ_u throughout the vertical, as expected. From the tropospheric boundary up to 500 K, the air ages relatively more quickly in the vertical than it does above 500 K. This change in the rate of aging in the vertical is more readily seen in the extratropical age Γ_d as compared to the tropical age Γ_u. The layer between 380 and 500 K can be interpreted as a diffusive boundary layer between the older air in the stratosphere and the much younger tropospheric air below (Neu and Plumb 1999). The air in the troposphere is relatively well mixed across latitudes. As most of the tropospheric air enters the stratosphere through the tropical pipe, the transition in the rate of aging is much less abrupt in the tropics than it is in the midlatitudes.

The meridional age difference $\Delta\Gamma$ increases from 380 up to 500 K (where it is the highest), subsequently decreasing above. This is shown more clearly in Fig. 3c, which compares $\Delta\Gamma$ (dashed) to the mean residence time $M/\mathcal{M}$ (solid). Both quantities have very similar structure throughout the vertical but the age difference is consistently less than the upper-bound $M/\mathcal{M}$ at all levels. The difference between the solid and dashed curves can be attributed to the diffusive diabatic flux of age, an additional vertical transport unaccounted for in our theory. The presence of diabatic diffusion (Sparling et al. 1997) and numerical diffusion of age reduces the age difference between the downwelling and upwelling regions, as they enhance the net vertical transport of age by diabatic circulation (Linz et al. 2016).

Figure 3d shows the inferred mixing flux $\mu_{\rm mix}$ (solid) and the net poleward flux $\mu_{\rm net}$ for the finite-volume core, derived from the vertical gradients of Γ_u. Both the fluxes rapidly decay in the vertical, reflecting reduction in the diabatic flux associated with decreasing density. The mixing flux is weaker than the net poleward flux at all isentropic levels, though the two are on the same order of magnitude in the lower stratosphere. The mixing flux all but dies away above 600 K. The unphysical negative values reflect a breakdown of the model associated with the fact that the tropical pipe (and extratropics) are not uniformly mixed, i.e., the age is not uniform along isentropes within the tropics/extratropics, as seen in Fig. 3a. The model assumes that air entrained from the tropics leaves with the mean age of tropical air; that air on the boundary of the tropical pipe is somewhat older than the mean causes us to underestimate the

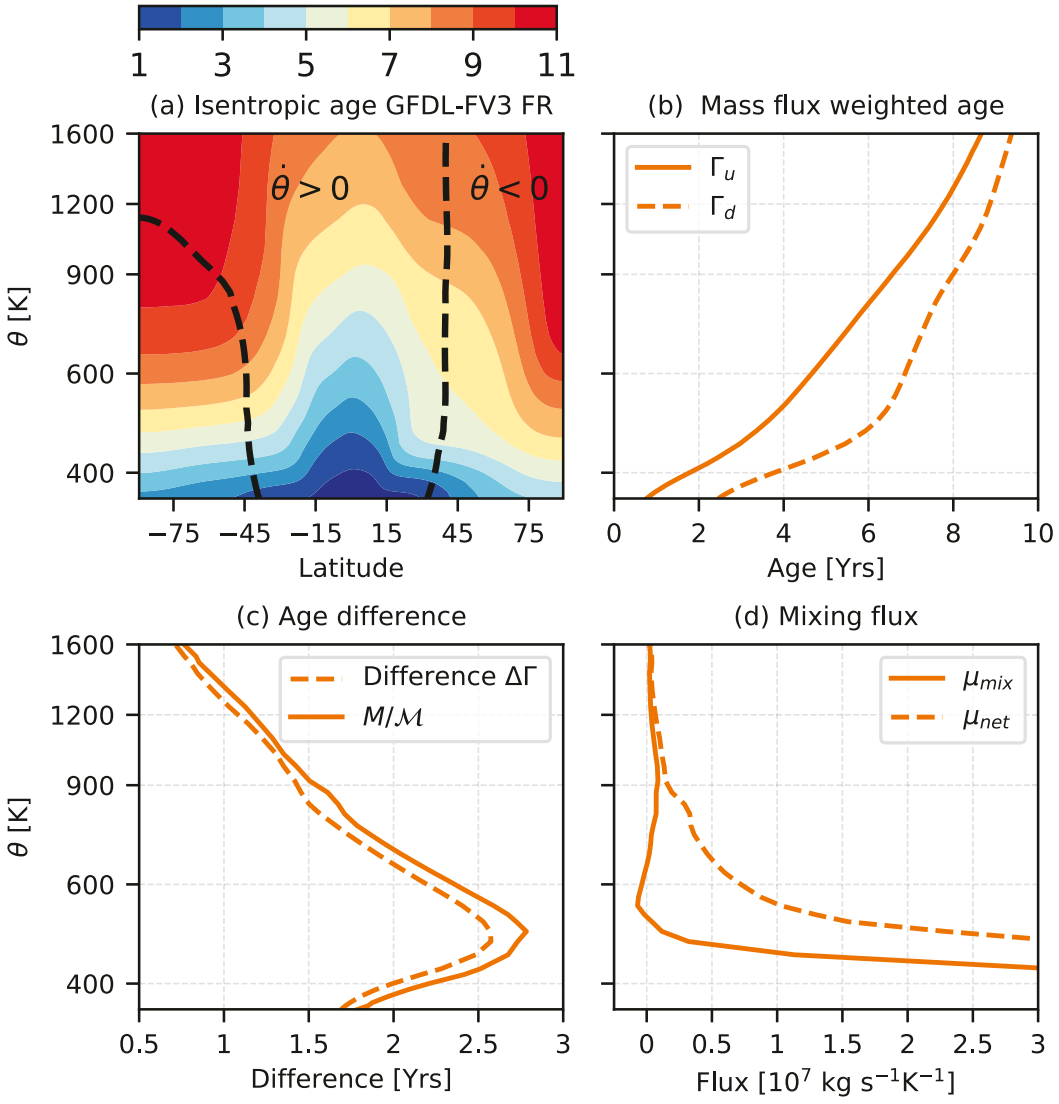

FIG. 3. Demonstration of the transport metrics introduced in section 3 for the free-running 1° L80 GFDL-FV3 model integration. (a) Zonal-mean age (in years) as a function of latitude and potential temperature. Dashed curves demarcate the upwelling ($\dot{\theta} > 0$) and downwelling ($\dot{\theta} < 0$) regions. (b) Mass-flux-weighted ages Γ_u (solid) and Γ_d (dashed). (c) The age difference $\Delta\Gamma$ (dashed) and the mean residence time $M/\mathcal{M}$. (d) The mixing flux μ_{mix} (solid) and the net poleward flux μ_{net} (dashed).

transport of age out of the pipe by the diabatic circulation. The mixing flux is computed as a residual, and so can become negative to balance the error. These negative (albeit near zero) values should be interpreted as very weak mixing, the amplitude of "negative" flux crudely quantifying uncertainty in our framework.

4. Mean age gradients: Quantifying vertical diffusion of age

We show the gross ages Γ_u (solid) and Γ_d (dashed) for the GFDL-FV3 (orange) and CAM-SE (green) in Fig. 4a. Since the FV3 core exhibits very similar age profiles for both the FR and SP test (on account of similar climatology), we only show the free-running case. For CAM-SE, however, we show results from both the FR and SP integrations.

At 380 K, Γ_u and Γ_d in both cores (and for both the FR and SP tests) are nearly identical, indicating that the age of air entering the tropical stratosphere is similar for all integrations (i.e., the models simulation of transport in the troposphere are fairly equivalent). The age profiles, however, diverge quickly with height. In CAM-SE, the FR test (dark green curves) exhibits substantially older air than the other two runs throughout the stratosphere in both the tropics (solid) and the extratropics (dashed). This difference from the FV3 core is reduced when the tropical winds in the CAM-SE core are prescribed in the SP test, albeit not entirely (light green curves with markers). We discuss these differences in detail in

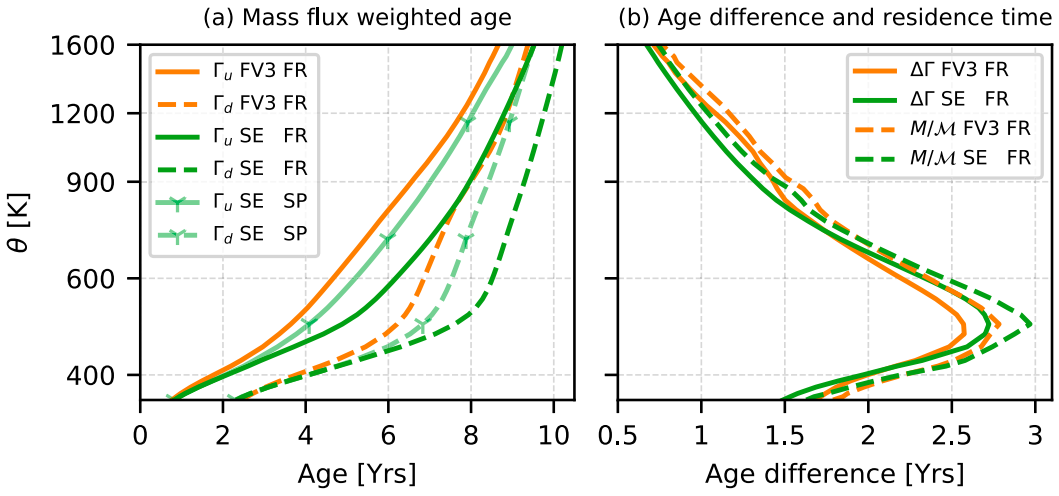

FIG. 4. (a) Gross upwelling and downwelling ages Γ_u (solid) and Γ_d (dashed) for GFDL-FV3 (orange) and the CAM-SE (green). The green curves with markers show the same quantities, but for the SP runs in CAM-SE. (b) Age difference $\Delta\Gamma$ and mean residence time $M/\mathcal{M}$ for the FR runs in the two models. The solid curves show the meridional age difference $\Delta\Gamma$ in the models and the dashed curves show the mean residence time $M/\mathcal{M}$ in the models.

section 5, but first focus on the age difference $\Delta\Gamma$ pictured in Fig. 4b, which is more closely tied to diabatic transport.

The age difference $\Delta\Gamma$ (solid curve) steadily increases from approximately 1.5–1.75 years at 380 K to 2.5–2.75 years at 500 K. Above 500 K, it steadily decreases in the vertical. This reflects the variation of the net mass flux $\mathcal{M}$ relative to the density of air (or equivalently, mass above the isentrope, M), i.e., variations in the upward velocity w, albeit inversely. There is a decrease in the upward velocity relative to density through the lower stratosphere, associated with the fall off in synoptic-scale wave forcing, and then a relative increase in the middle to upper stratosphere, associated with wave forcing from planetary waves above. In particular, between 400 and 750 K, the age difference (and residence times) are different between the FV3 and SE integrations, with SE exhibiting a slightly larger $\Delta\Gamma$ and $M/\mathcal{M}$ than FV3. This reflects a weaker diabatic circulation ($\mathcal{M}$) in SE, not any material differences in M.

The dashed curves, which show the mean residence time of air above a given isentrope, have a very similar vertical structure to $\Delta\Gamma$, but with slightly higher magnitude. The separation between the solid and corresponding dashed curves, i.e., the residual in Eq. (2), represents the contribution to the age difference due to both diabatic dispersion (a physical phenomenon) and any numerical diffusive fluxes in the models. [In the real atmosphere, there are also turbulent diffusive fluxes, which primarily arise from breaking internal gravity waves (Dewan 1981), but this is not represented in the dynamical cores.] Diabatic dispersion is weak in the tropics, where the gradient in diabatic heating is small, though it is potentially stronger in the extratropics where there are meridional gradient in diabatic heating (Sparling et al. 1997; Plumb 2007). We also expect numerical diffusion to be weaker in the tropics (where isentropes line up with pressure) than in the extratropics (where they do not), as the vertical diffusion at higher latitudes will also include a projection of the horizontal model diffusion on to the cross-isentropic diffusive fluxes. We cannot separate diffusion from dispersion, but the strong sensitivity to vertical resolution (see below) suggests that numerical diffusion may be dominant.

These effects are quantified in the models by comparing the observed age difference $\Delta\Gamma$ with that implied by advection alone, the mean residence time $M/\mathcal{M}$, as shown in Fig. 4b. The diffusive flux of age, proportional to the negative of vertical age gradient, $-\partial_\theta\Gamma$, is always negative (acting downward) because the age monotonically increases with height, and hence reduces the meridional age difference relative to the mean residence time computed from advection alone. In the presence of vertical diffusion, the mean residence time on the right-hand side of Eq. (2) is determined as the ratio of the total mass above the isentrope M and the total mass flux $\mathcal{M} + \mathcal{F}_D$, where $\mathcal{F}_D$ is the diffusive mass flux. As a result, the horizontally averaged vertical diffusive mass flux can be estimated as the quantity $\mathcal{F}_D = M/\Delta\Gamma - \mathcal{M}$ by rearranging the terms. The net vertical diffusive age flux can then be obtained by multiplying the diffusive mass flux $\mathcal{F}_D$ with the average age between the two regions, $(\Gamma_u + \Gamma_d)/2$. We find it most effective to compare this diffuse flux of age to the total advective flux of age across the isentrope, $\mathcal{M}_u\Gamma_u + \mathcal{M}_d\Gamma_d$. Thus we plot the ratio

$$\mathcal{D} = \frac{\frac{\Gamma_u + \Gamma_d}{2}\left(\frac{M}{\Delta\Gamma} - \mathcal{M}\right)}{\mathcal{M}_u\Gamma_u + \mathcal{M}_d\Gamma_d}. \tag{4}$$

Figure 5a shows the ratio $\mathcal{D}$ for 1° FV3 (orange) and SE (green) cores at three different vertical resolutions. For simplicity, we only show the SP integrations. Our analysis shows that in the lower stratosphere, the diffusive tracer flux of age accounts for up to 10% of the resolved tracer flux. Moreover, this fraction rapidly decreases with height, more so for the 40-level runs (dotted curves). This finding is consistent with the diffusive boundary layer identified in Fig. 4, which indicates that the

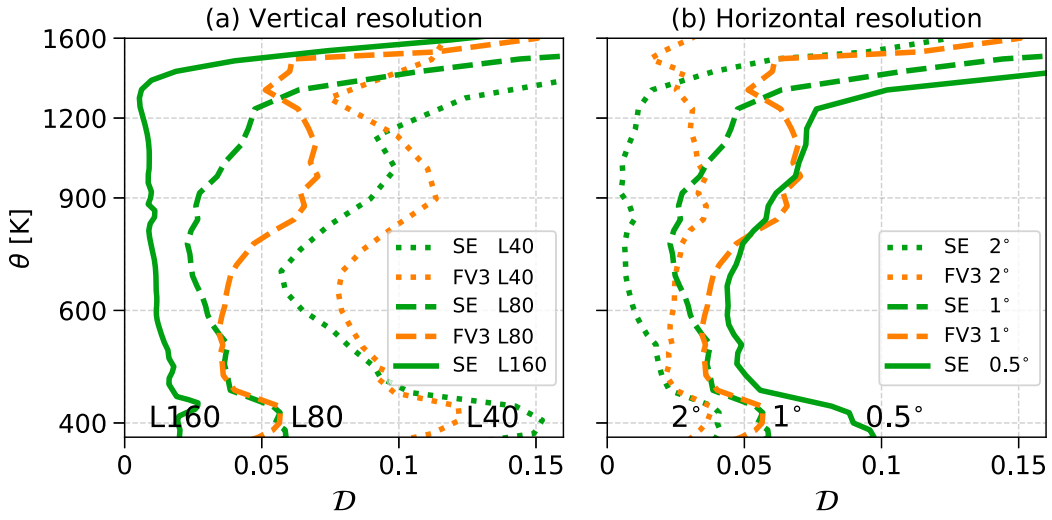

FIG. 5. Effect of changing (a) the vertical resolution and (b) the horizontal resolution on the diffusive fluxes in the FV3 core (orange) and the SE core (green). The nondimensional ratio $\mathcal{D}$, defined in Eq. (4), is plotted: the ratio of the diffusive flux of age relative to the total advective flux of age. The dotted curves, dashed curves, and solid curves in (a) represent SP runs with 40, 80, and 160 levels in the vertical at 1° horizontal resolution. The dotted curves, dashed curves, and solid curves in (b) represent SP runs with 2°, 1°, and 0.5° resolution in the horizontal resolution with 80 levels in the vertical. All the curves shown correspond to the specified tropics runs for the two models.

shapes for Γ_u and Γ_d in the lower stratosphere are largely determined by the diffusive stratospheric–tropospheric exchange of age.

As the models' vertical resolution is doubled from 40 to 80 levels (and subsequently from 80 to 160 levels), a rapid decrease in the contribution from the numerical vertical fluxes is observed. For integrations with 160 levels in the vertical, the contribution from vertical diffusion is practically insignificant enough to not affect vertical transport. Figure 5a highlights the importance of sufficiently high vertical resolution in idealized and comprehensive stratosphere-resolving climate models used to study UTLS and stratospheric transport. Moreover, it indicates that the upper tropospheric and lower-stratospheric region is relatively more sensitive to the effects and consequences of model diffusion, as compared to the middle and upper stratosphere.

Figure 5b shows the same diffusive-to-advective age flux ratio, but for 80 vertical level FV3 and SE cores integrated at three different horizontal resolutions. We observe the seemingly opposite dependence on resolution: refining the *horizontal* grid increases *vertical* diffusion. As the horizontal resolution is increased from 2° to 1° (and subsequently from 1° to 0.5° for CAM-SE), the diffusive tracer flux increases in both the FV3 and SE core. This dependence of the vertical diffusion on aspect ratio suggests that the net vertical diffusion may be effected by quasi-horizontal diffusion across isentropic surfaces near the grid scale. Allowing additional horizontal motions increases effective diffusion if the vertical resolution is not increased at the same time to properly resolve the finest allowed scales. This perverse effect highlights the importance of considering the aspect ratio when refining the grid of an atmospheric simulation, as highlighted by Lindzen and Fox-Rabinovitz (1989).

While vertical diffusion varies with resolution in both cores, the effect is qualitatively the same, and quantitatively similar, particularly in the lower stratosphere: vertical diffusion does not explain the difference in tracer transport between the two cores. We thus turn to differences in the diabatic circulation and adiabatic mixing.

5. Quantifying the impact of mixing and the diabatic circulation

Following Linz et al. (2016, 2021) we connect the full three-dimensional transport in the dynamical cores to the one-dimensional TLP model of Neu and Plumb (1999). Our goal is to construct an analog to the TLP model from three-dimensional dynamical core integrations. The key is to allow for vertical variations in the TLP model parameters, which are listed in Table 2.

We first compute the optimal value for each parameter based on the dynamical core output. Allowing the coefficients to vary in the vertical precludes closed form, analytic solutions, so we numerically solve the equations using a second-order implicit Crank–Nicolson temporal integrator to obtain the mean tropical and midlatitude ages, Γ_T and Γ_M. After verifying these solutions against the corresponding ages in dynamical cores, Γ_u and Γ_d, we can characterize the importance of differences in mixing, circulation, and other factors by systematically perturbing the parameters.

The vertical coordinate of the TLP is switched from height to entropy (potential temperature θ) and all the isentropic averages are obtained using the binning technique of Yamada and Pauluis (2015); see the online supplementary material for details. The average vertical velocity $\dot{\theta}$ is computed over the upwelling and

TABLE 2. (middle) The original leaky pipe parameters and (right) the corresponding climate model analogs used to compute the vertically varying parameters in the leaky pipe model. The diffusion coefficient in the tropics was tuned to match the age from the models and from the solvers and is model dependent. Both models use an identical value of $K_M = 1.5 \times 10^{-2}\,\mathrm{m^2\,s^{-1}}$ for the midlatitude diffusion coefficient. For GFDL-FV3, $K_T = K_M = 1.5 \times 10^{-2}\,\mathrm{m^2\,s^{-1}}$ and for CAM-SE, $K_T = 0.5\,K_M = 0.75 \times 10^{-2}\,\mathrm{m^2\,s^{-1}}$. It is reassuring that the values of K_T and K_M obtained in our analysis are similar and that $K_M \geq K_T$, since horizontal diffusion can project onto the vertical in the extratropics. The magnitude of K_M used by Neu and Plumb (1999) is notably higher than that obtained in our study; however, Neu and Plumb (1999) noted that their choice of K_M was purposely overestimated to assess the effects of vanishing diffusion on age.

Physical quantity	Neu and Plumb (1999)	Model analog
Scale height (H)	7 km	6.5 km
Vertical velocity (W_T)	$0.3\times 10^{-3}\,\mathrm{m\,s^{-1}}$	$\dot{\theta}\left\lvert\frac{d\theta}{dp}\right\rvert = \frac{\int_u \rho_\theta \dot{\theta}\, dA}{\int_u \rho_\theta\, dA}\left\lvert\frac{d\theta}{dp}\right\rvert$
Mass distribution (α)	0.5	$\frac{\sigma_u}{\sigma_u + \sigma_d}$
Entrainment ratio (λ)	$-\frac{\alpha W_T}{H}$	$-\frac{\partial_\theta \mathcal{M}_u}{\sigma_d}$
Mixing (ε)	$\in (0, 1)$	$\frac{\mu_{\mathrm{mix}}}{\mu_{\mathrm{net}}}$
Upwelling age (Γ_T)	0 ($z = 0$)	Γ_u ($\theta = 380$ K)
Downwelling age (Γ_M)	0 ($z = 0$)	Γ_d ($\theta = 380$ K)
Vertical diffusion, tropics (K_T)	$1 \times 10^{-2}\,\mathrm{m^2\,s^{-1}}$	See caption
Vertical diffusion, midlatitudes (K_M)	$0.5\,\mathrm{m^2\,s^{-1}}$	$1.5 \times 10^{-2}\,\mathrm{m^2\,s^{-1}}$

downwelling regions, weighted by mass. Similarly, the isentropic densities σ_u and σ_d are computed by horizontally integrating the isentropic density ρ_θ over the upwelling and downwelling partitions, respectively. The mass distribution α between the tropical pipe and the extratropics is estimated as the ratio $\alpha = \sigma_u/(\sigma_u + \sigma_d)$. The scale height in our model stratosphere does not significantly deviate from 6.5 km and so $H = 6.5$ km is chosen.

The mass-flux-weighted ages Γ_u and Γ_d, analogous to the tropical and midlatitude ages Γ_T and Γ_M of the original Neu and Plumb model, were already described. The mixing flux from the midlatitudes into the tropics, μ_{mix}, is estimated using Eq. (3) (and illustrated in Fig. 6), while the net mass flux μ_{net} is computed as the vertical derivative of diabatic mass flux $\mathcal{M}(\theta)$. The two quantities are used to compute the TLP parameter ε, the mixing efficiency. Finally, the vertical diffusion parameters K_T and K_M are tuned to ensure that the model data estimates of Γ_u and Γ_d are in agreement with Γ_T and Γ_M obtained by numerically integrating the TLP equations. We find that for the midlatitudes, a fixed diffusion coefficient of $K_M = 1.5 \times 10^{-2}\,\mathrm{m^2\,s^{-1}}$ provides the best fit. The optimal tropical diffusion coefficient K_T, however, slightly varied between the dynamical cores: $1.5 \times 10^{-2}\,\mathrm{m^2\,s^{-1}}$ for FV3, but $0.75 \times 10^{-2}\,\mathrm{m^2\,s^{-1}}$ for SE.

As discussed in section 3c, the three-dimensional circulation cannot be perfectly mapped onto the 1D framework. A key assumption of Neu and Plumb (1999) is that fast mixing within the tropical pipe and the midlatitudes completely homogenizes the age in each region; all latitudes in the tropics (midlatitudes) are assumed to have the same age Γ_T (Γ_M) and it does not matter from where air is entrained or mixed between the layers.

This is not the case for the simulated stratosphere in our idealized climate models. Figure 3a illustrates meridional variations of age in the two regions. Air entrained out of the tropics (and out of the midlatitudes) will be older than the mean tropical age Γ_u (and younger than the mean midlatitude age Γ_d) leading to a systematic positive biased flux of age between the layers. As a result, the mixing flux μ_{mix} will be systematically negatively biased, to the point that it can become unphysically negative. While this reflects a weakness in the model, it does not affect our ability to reconstruct the age in the two integrations and assess the impact of varying all our parameters.

Several past studies have explored the TLP in the context of models and observations (Ray et al. 2010; Garny et al. 2014; Ray et al. 2016; Kolonjari et al. 2018). Ray et al. (2010), for example, applied the TLP to the observed ozone profiles and inferred mean age profiles to connect the multidecadal ozone variability to changes in mean circulation and isentropic mixing. Similarly, Garny et al. (2014) used a linear fit to infer a (vertically invariant) value for the mixing efficiency ε in climate model simulations. They did not account for vertical variations in transport, which can prevent a direct comparison between the model age and the age in the TLP.

a. *Mixing fluxes versus diabatic fluxes*

The mixing flux μ_{mix} is plotted alongside the net poleward flux μ_{net} for the free-running GFDL-FV3 and CAM-SE dynamical cores at 1° L80 resolution in Fig. 6a. The two models exhibit very similar net meridional fluxes μ_{net} (solid curves). The μ_{net} is equal to the vertical convergence of the diabatic mass flux $\mathcal{M}$, by mass continuity; the two free-running integrations thus exhibit a fairly comparable diabatic circulation, at least in terms of the gross overturning. Minor differences in μ_{net} do affect the transport, but are swamped in the difference in the mixing flux between the two free-running integrations.

The mixing fluxes μ_{mix} (dashed lines) differ greatly between 400 and 600 K. At 500 K, for instance, $\mu_{\mathrm{mix}} = 0.5 \times 10^7\,\mathrm{kg\,s^{-1}K^{-1}}$

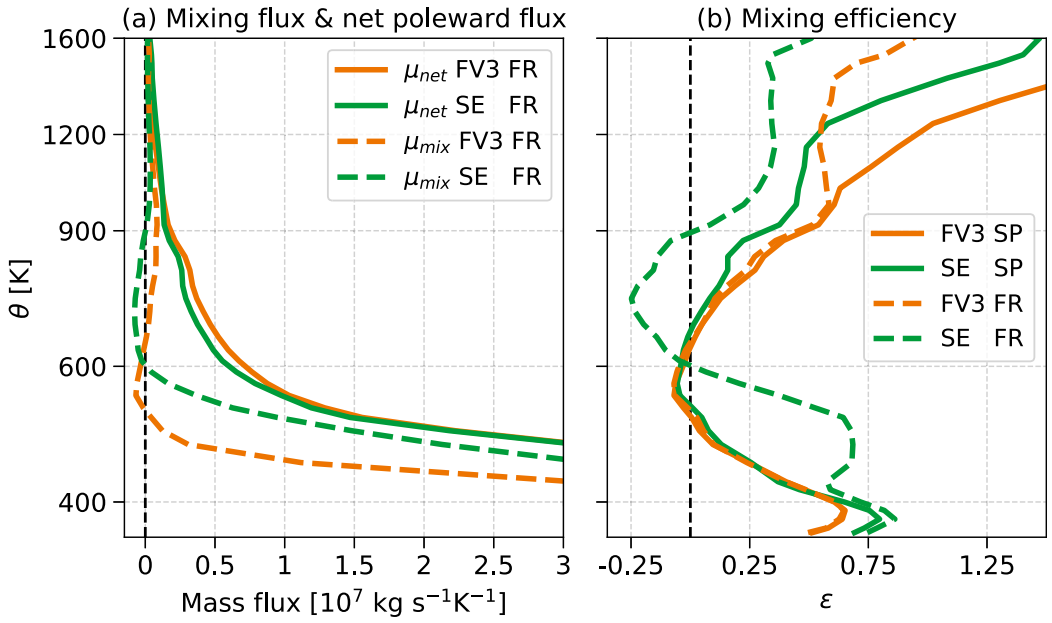

FIG. 6. (a) The mixing flux μ_{mix} (in $10^7\,\text{kg}\,\text{s}^{-1}\,\text{K}^{-1}$) and (b) the mixing efficiency ε for the GFDL-FV3 core (orange) and CAM-SE core (green). The net poleward flux μ_{net} in (a) is shown in solid curves and the mixing flux μ_{mix} is shown in dashed curves. The dashed (solid) curves in (b) show the mixing efficiency for the free-running (specified) integrations.

for the FV3 core, compared to $2.5 \times 10^7\,\text{kg}\,\text{s}^{-1}\,\text{K}^{-1}$ for the SE core: at this level, 5 times the mass is being exchanged between the tropics and midlatitudes in the SE core, even though the net flux from tropics to the extratropics is equivalent. The mixing flux quickly decays to near zero in the midstratosphere, but this occurs circa 550 K in the FV3 core versus circa 600 K in the SE core. Figure 6a only shows the FR integrations; the mass fluxes obtained for the SP integrations for the two cores are nearly identical to those of the FV3-FR integration, as can be inferred from the similarity of the mixing efficiency in Fig. 6b.

Given the exponential fall off in the mass flux (associated with the exponential fall off in density), it is more illustrative to plot the mixing efficiency $\varepsilon = \mu_{\text{mix}}/\mu_{\text{net}}$, shown in Fig. 6b. The mixing efficiency measures the "leakiness" of the tropical pipe. In the limit $\varepsilon \to 0$, the tropics and midlatitudes are entirely isolated, apart from the net poleward flow of mass associated with the diabatic circulation, as in the model of Plumb (1996). In the $\varepsilon \to \infty$ limit, explored by Plumb and Ko (1992), the mixing time scale entirely dominates the diabatic flux.

Focusing first on free-running integrations (dashed curves in Fig. 6b), we see that the mixing flux peaks relative to the diabatic flux in the lowermost stratosphere ($\approx$390 K) in both models. The key difference between them is in how the mixing falls off above. The peak in ε is associated with enhanced mixing due to synoptic-scale wave breaking at the top of the subtropical jets. Above, the mixing efficiency steadily decreases up to 550–600 K, where it ultimately vanishes. The reduction in mixing in the middle and upper stratosphere is consistent with the findings of Haynes and Shuckburgh (2000), who quantified mixing using the effective diffusivity proposed in Nakamura (1996). Their computations also indicated a very low eddy diffusivity in the tropics above 600 K.

The solid curves in Fig. 6b show the mixing efficiency in the two integrations where the tropical winds were constrained to an easterly profile. There is little difference in mixing between the SP and FR in FV3, consistent with the fact that both models exhibit similar winds in the tropics. The difference between the FR and SP integrations of the SE core indicate that the tropical wind structure strongly controls the rate of mixing. The transition from a climatological state with westerly to easterly winds centered about 600 K (Fig. 1b vs Fig. 1e) drops the ceiling of the high mixing rates, bringing the spectral element core with specified winds into a state similar to the finite-volume-based core. Westerly winds shift the critical lines deeper into the tropics, allowing deeper penetration of Rossby waves into the tropics, leading to enhanced mixing. Enhanced mixing is associated with the westerly phase of the QBO (Plumb and Bell 1982); here the locked westerlies lead to a climatological increase in mixing. As detailed in G20, the specified easterly wind profile was chosen to match the winds observed in all integrations with low vertical resolution; this choice is not meant to imply that easterlies are the "correct" profile, and one could alternatively increase the mixing in all the models by specifying a westerly profile.

As discussed in section 3c, the unphysically negative values of the mixing flux, most pronounced in the free-running integration with the spectral element core, reflect the limitation of our method. High mixing between 400 and 600 K in the SE integration leads to a situation where the age is not well mixed across the tropical pipe. The age distribution is more sharply peaked in the tropics in Fig. 1b relative to the other integrations, exacerbating this error. We believe that the

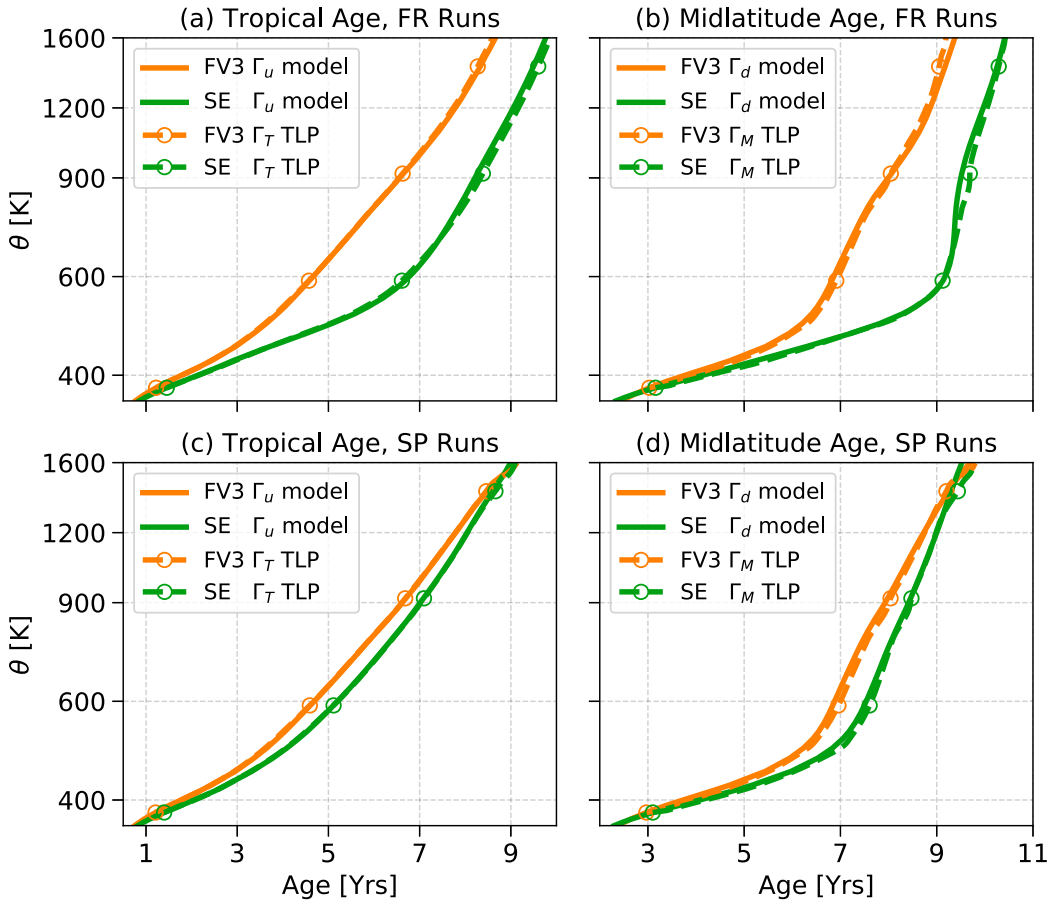

FIG. 7. The match between the (a),(c) tropical and (b),(d) midlatitude age in our leaky pipe emulator and the original model integrations. The solid orange (green) curves show the gross ages Γ_u and Γ_d, obtained by mass-flux-weighted averaging output from the GFDL-FV3 (core) and the dashed orange (green) curves show the corresponding ages Γ_T and Γ_M, obtained from the TLP emulator when forced with the parameters fit from model output.

correct interpretation is that mixing is essentially zero above 600 K, at least until the upper stratosphere above 900 K.

b. *The global impact of enhanced mixing in the lower stratosphere in the FR integrations*

The leaky pipe emulator allows us to diagnose the impact of the difference in mixing between 400 and 600 K on tracer transport throughout the stratosphere.

Using the parameters computed from the 3D model integrations (tabulated in Table 2), we first establish that one can reconstruct the age in the tropics and extratropics, Γ_u and Γ_d, with our numerical solver of the variable coefficient TLP model. As illustrated in Fig. 7, there is a good match with Γ_T and Γ_M, respectively; here we have kept the notation of Neu and Plumb (1999) to distinguish the ages from the dynamical core integrations and the TLP emulator. The fit is so good for the tropical pipe in the SP integrations that Γ_u and Γ_T cannot be distinguished. This is not a surprising result—the parameters were fit with knowledge of the ages Γ_u and Γ_d—but establishes that the emulator works well with fixed diffusivity coefficients in the vertical.

To assess the contribution of each transport process to the spread in age distributions in the model integrations, we take the following approach. Starting with the parameters obtained for the FV3 core, we perturb the parameters, one at a time (or in related groups), computing the impact of each process on the mean age of the tropics and extratropics. Figure 8 shows the results for the free-running integrations. For example, the orange curve in Fig. 8a shows Γ_T in the TLP emulator, our fit to the mass weighted age Γ_u of the FR integrations of the FV3 core. The blue dashed curve shows the TLP emulator tropical age when the mixing efficiency ε is changed from its value in the FV3 fit to the SE fit, all other parameters left the same. The difference in mixing alone explains the bulk of the difference between the two integrations, as shown explicitly in Fig. 8b.

Before commenting on each process, we first establish that the results are sufficiently linear to separate them. The purple dotted curve in Fig. 8b shows the sum of differences associated with each process, compared to the total difference between Γ_T fit for the SE and FV3 models in the green

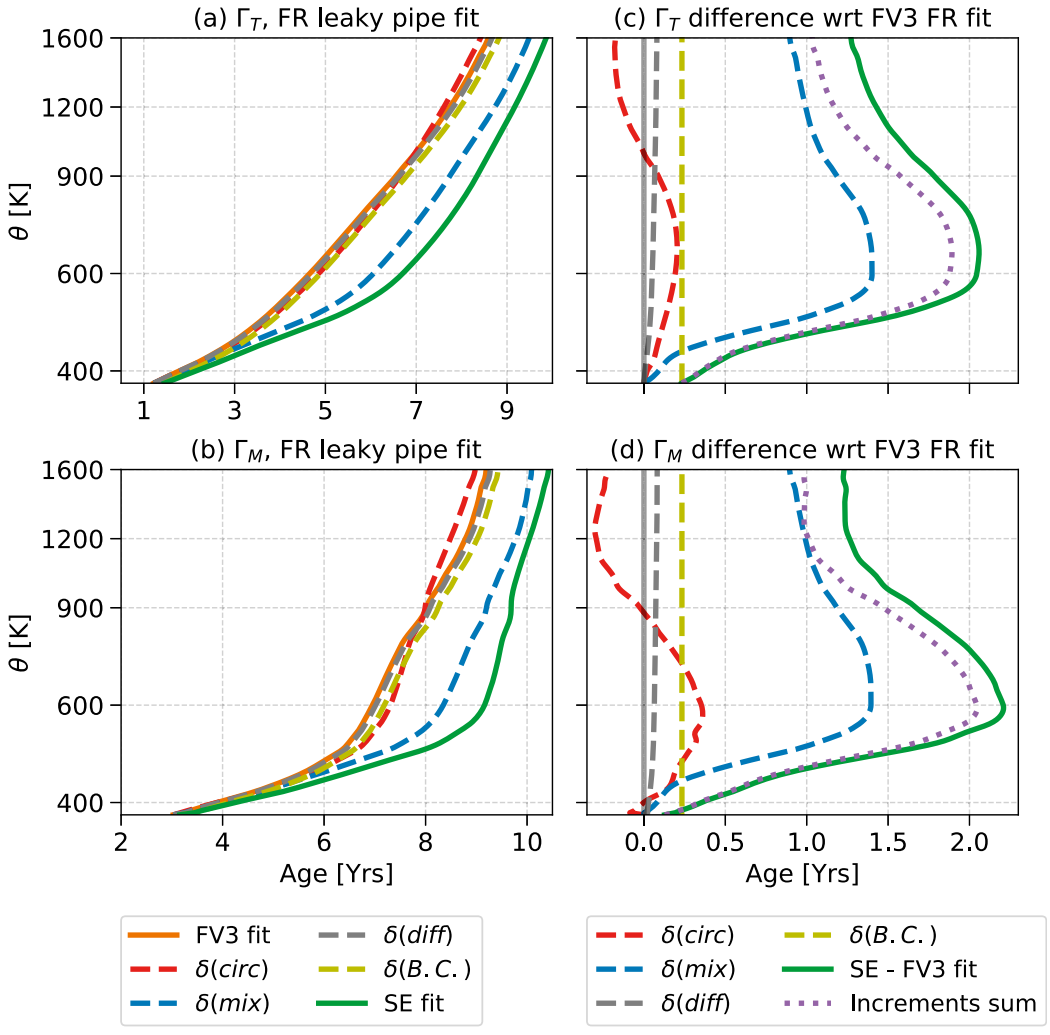

FIG. 8. Perturbation analysis using the TLP emulator. (a) The tropical age fit Γ_T and (b) the midlatitude age fit Γ_M obtained from the numerically integrated TLP. Orange and green curves show the ages obtained using GFDL-FV3 and CAM-SE parameters, while the red, blue, gray, and yellow curves show the age obtained when parameters quantifying the diabatic circulation strength, isentropic mixing, diffusion, and 380 K boundary conditions from CAM-SE are *individually* imposed onto the GFDL-FV3 core (i.e., only one parameter is changed, and the others are left at the values of the GFDL-FV3 fit). (c),(d) The difference obtained (with respect to the FV3 fit) for each of the increments corresponding to the same color as in (a) and (b). The green curve shows the total age difference between the green and orange curves in (a) and (b), i.e., the difference between the SE and FV3 age. The dotted violet curve in (c) and (d) is the sum of individual age increments obtained by substituting SE core's parameters; its similarity to the total difference (green) quantifies nonlinearity between the perturbations.

curve. For the FR integrations, the sum of the parts is always very good (≈95% of the total) up to 1000 K. Above this altitude, nonlinearity starts to matter (i.e., the change in age associated with mixing is exacerbated by differences in other processes). For the SP integrations, where differences are smaller, the result is extremely linear.

Four parameter groups characterize the key processes that govern stratospheric transport:

(i) the diabatic circulation: μ_{net} and α (which characterizes the geometry of the flow);
(ii) isentropic mixing: ε;
(iii) numerical diffusion; and
(iv) the input of age at the tropopause: the age at $\theta = 380$ K.

The impact of switching these parameters from the optimal FV3 value to those based on SE are shown in Fig. 8. The differences associated with the diabatic circulation (in red) and boundary conditions (in yellow) do not explain the tropical and midlatitude age differences between the FV3 and SE free-running integrations. A striking jump in age, however, occurs when the SE mixing profile is imposed onto FV3. The difference in mixing (blue curves in Fig. 8) accounts for almost three-quarters of the net age difference between FV3 and SE free-running integrations (green curves).

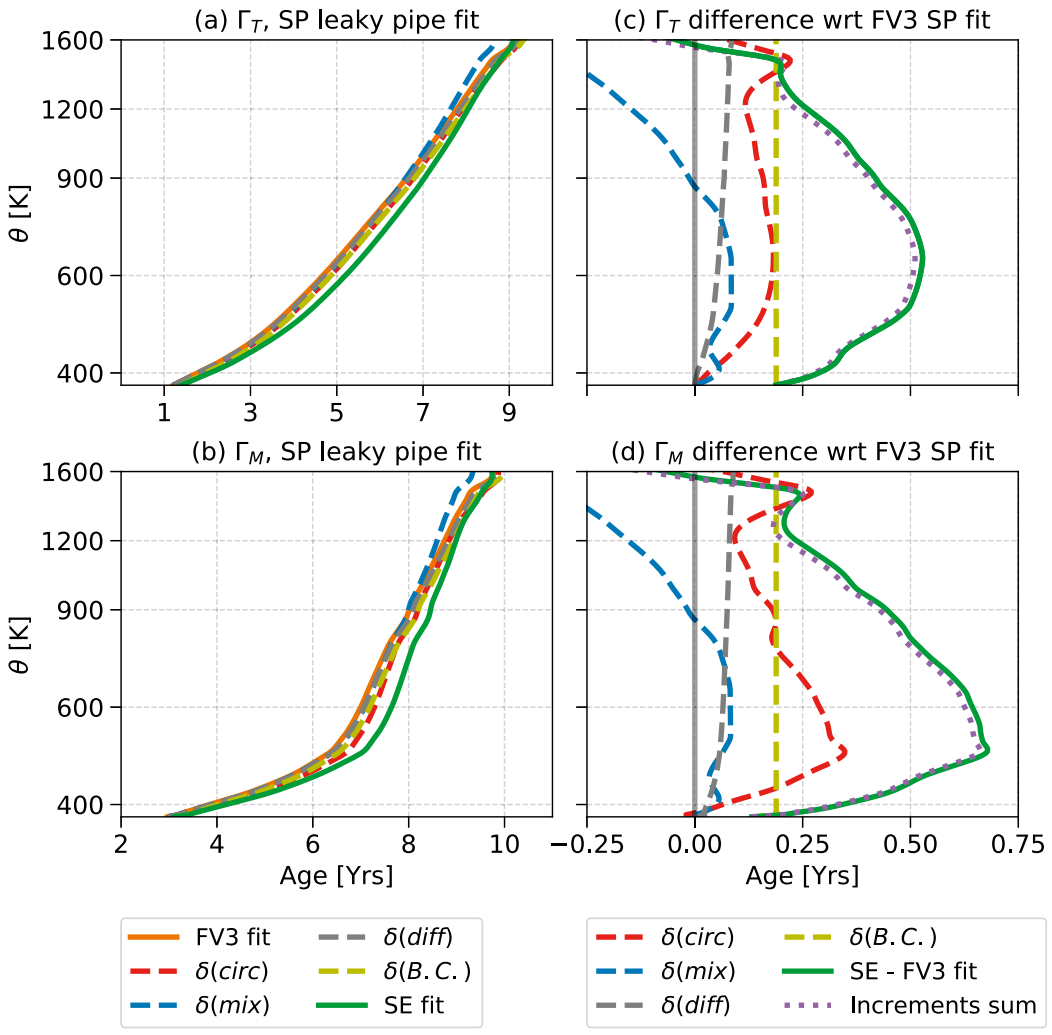

FIG. 9. As in Fig. 8, but for the specified tropics (SP) model test integrations. Note the smaller x range for (c) and (d) as compared to Figs. 8c and 8d; transport in the two cores is much more similar in the SP integrations compared to the FR integrations.

The contribution due to difference in diabatic circulation strength is small, not greater than 10% in the middle stratosphere, though weaker overturning in the lower stratosphere in the SE model does amplify the aging associated with mixing. The contribution from differences in boundary conditions (i.e., age transport in the troposphere) is relatively small as well, not accounting for more than 10% of the age difference. The minor difference left unaccounted by parameter groups (i)–(iii) is the difference in model diffusion between the CAM-SE and GFDL-FV3 cores. Nonlinear effects matter in the upper stratosphere, as differences in age due to mixing and circulation lead to an overall greater difference between the cores.

c. *The importance of circulation differences in the SP integrations*

We repeated the TLP analysis for the SP runs in Fig. 9. For the integrations with similar tropical wind climatologies, the net difference in age is much smaller as compared to the FR runs. Figure 9 suggests that most of the age differences among the SP runs are caused by differences in the diabatic circulation (dashed red) and boundary conditions associated with transport through the troposphere (dashed yellow) between the cores. Differences in boundary conditions provides a straight offset of 0.2 years (up to 50%) in both the tropics and the midlatitudes. Further, differences in diabatic circulation accounts for almost 50% of the total age differences in the middle stratosphere. Differences due to mixing are relatively small up to 900 K. Enhanced mixing in the SE core leads to an aging of air in the lower stratosphere, while slightly weaker mixing in SE (compared to the FV3 core; solid curves in Fig. 6b) above ~800 K results in a slight reduction in age in the upper stratosphere (dashed blue curve in Fig. 9).

Even though the relative contribution of diabatic velocity differences is lower for the FR runs than for the SP runs, the absolute contribution remains quite similar. More simply put, for both the FR and SP runs, differences in diabatic

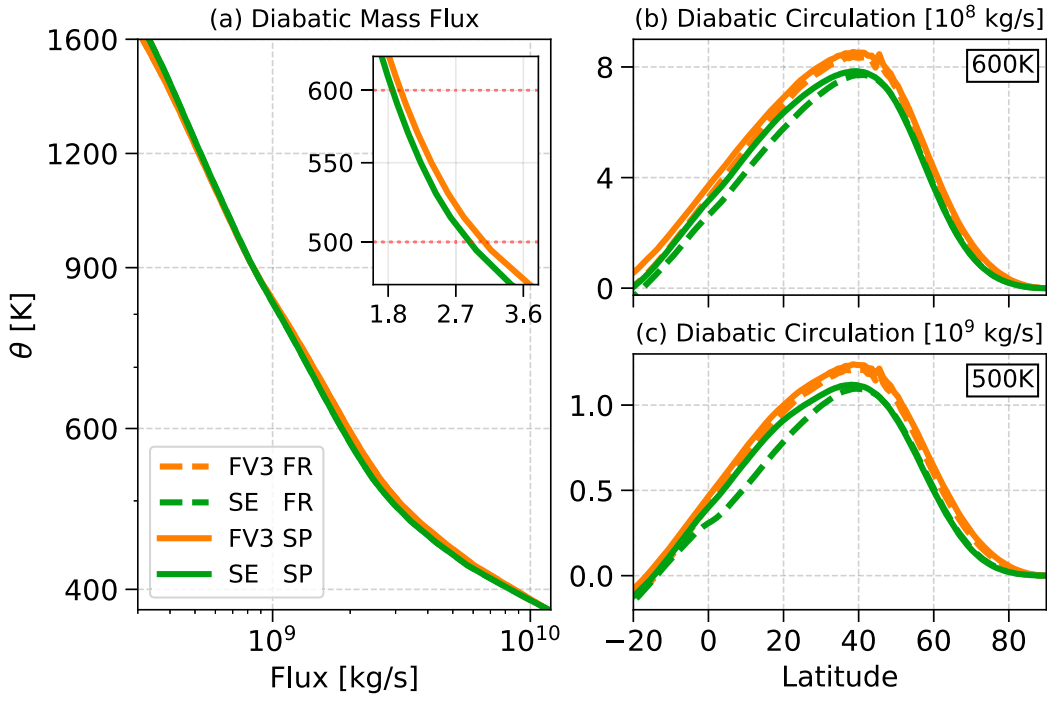

FIG. 10. (a) Horizontally integrated diabatic mass flux (in kg s^{-1}) in the upwelling region, $\mathcal{M}_u$, for the 1° L80 SP runs for the GFDL-FV3 (orange) and the CAM-SE (green) cores. The inset shows a close-up of the flux between 450 and 650 K. (b),(c) The full diabatic streamfunction $\psi(\phi, \theta)$ for the FR runs (dashed) and SP runs (solid) at (b) $\theta = 600$ K and (c) $\theta = 500$ K.

circulation accounts for as much as 0.25 years (0.4 years) of the net difference in Γ_T (Γ_M). This suggest that differences in the diabatic circulation are unrelated to differences in the tropical wind climatology and unaffected by the mild damping used to constrain the winds. The diabatic circulation $\mathcal{M}$ for the 1° L80 integrations for the two models is compared in Fig. 10. The FV3 core consistently exhibits a stronger circulation, particularly in the middle stratosphere (450–700 K).

To compare the meridional structure of the diabatic circulation among models following G20, we compute the diabatic streamfunction (Townsend and Johnson 1985; Pauluis et al. 2009), shown in Figs. 10b and 10c. As in Pauluis et al. (2009), the time-averaged Eulerian streamfunction ψ in isentropic coordinates can be expressed as an integral of the meridional mass flux,

$$\psi(\phi, \theta_0) = \frac{R\cos\phi}{g(T_2 - T_1)} \int_{T_1}^{T_2} \int_0^{2\pi} \int_0^{p_s} vH[\theta(\lambda, \phi, p, t) - \theta_0]\, dp\, d\lambda\, dt, \tag{5}$$

where the time interval is (T_1, T_2), v is the meridional velocity, H is the Heaviside function, which equals 1 for nonnegative values and 0 otherwise, λ is the longitude, ϕ is the latitude, R is the radius of Earth, and g is the gravitational acceleration. The time averaged diabatic circulation strength is obtained by averaging the streamfunction over the last 8400 model days.

For the free-running integrations, the key difference between the models is in the tropics, where westerlies induce a secondary mean meridional circulation (Plumb and Bell 1982). The presence of this mean meridional circulation can be seen in Fig. 10c as a shoulder in the green dashed curve at 500 K between the equator and 30°N, which is not observed at 600 K (Fig. 10b). When the winds in the tropics are constrained, the shoulder in the circulation for CAM-SE at 500 K disappears and a strengthened circulation for the core is noticed at both 500 and 600 K. For the FV3 core (which has similar tropical winds for both FR and SP runs) the circulation strength is quite similar.

The two state-of-the-art models still exhibit differences in the circulation strength in the middle stratosphere, however, despite being identically forced. While these differences are overshadowed by major differences in mixing in the FR runs, for the SP integrations, they account for up to half of the overall age difference between the models.

6. Conclusions

We have assessed how the formulation of an atmospheric model's dynamical core impacts stratospheric transport using the theory of age transport in isentropic coordinates developed by Linz et al. (2016, 2021). We compare two state-of-the-art dynamical cores, GFDL's Finite Volume 3 (FV3) and the Community Atmosphere Model's Spectral Element (CAM-SE), which employ very different underlying numerical methods to

solve the primitive equations (section 2). Both models were run at comparable resolutions and driven with identical diabatic and tracer forcings, as prescribed in the transport benchmark tests established by G20. Despite the carefully prescribed test environment, the two cores diverge substantially in their representation of stratospheric transport, particularly in a "free-running" configuration where the models produce a very different representation of the zonal winds in the tropical stratosphere, as highlighted in Fig. 1 and detailed by G20. Variations in transport can arise from differences in the simulation of the explicitly resolved circulation (both the slow overturning circulation, or diabatic circulation, and the comparatively faster adiabatic mixing of air along isentropes), and implicit differences in transport by the numerical schemes (trace gas representation, gridscale diffusion, and other errors). An analysis of the age budget in isentropic coordinates and a vertically varying formulation of the tropical leaky pipe model (TLP; Neu and Plumb 1999) allowed us to diagnose the individual impact of each factor on the circulation.

As the transport is sensitive to vertical resolution, particularly in CAM-SE, we first quantified the contribution of numerical diffusion in section 4. While the contribution from diffusion is tied to vertical resolution (Fig. 5), it does not explain the gross differences between the models. At moderate vertical resolution (40 vertical levels), the diffusive tracer flux is on the order of 10% of the resolved diabatic fluxes in both models. Its contribution rapidly decreases with increasing vertical resolution, becoming negligible in 160 level runs. Since most comprehensive climate models still employ less than 40 levels in the vertical, our findings highlight the importance of vertical resolution for studying stratospheric transport. It is differences in the explicitly resolved tracer transport, however, that dominate differences between GFDL-FV3 and CAM-SE in the benchmark experiments.

As reviewed in section 3, the vertical gradient of the gross upwelling age Γ_u allows one to quantify the adiabatic mixing flux μ_{mix}, i.e., the transport of tracer between the tropics and extratropics along isentropic surfaces by breaking waves. The key to relating age in a three-dimensional model to the TLP is to weight the flow by the mass flux when computing the average age Γ_u. Equation (3) relates the vertical age gradient to the aging purely by diabatic advection versus "aging by mixing," allowing us to estimate the mixing flux in the two models as a residual. With the vertically varying TLP emulator, we then quantified the impact of differences in the mean diabatic circulation versus differences in mixing.

For the free-running (FR) integrations, differences in transport are dominated by mixing. The mixing flux varied considerably in the middle stratosphere (450–650 K): CAM-SE exchanges up to 5 times more air between the midlatitudes and the tropics compared to GFDL-FV3 (Fig. 6). This enhanced mixing accounts for up to 75% of the total observed age differences among the FR model runs (Fig. 8). The specified tropical wind test integrations (SP test) establish that it is the zonal wind profile in the tropics that controls this difference in mixing. The formation of westerly jets in CAM-SE allows deeper penetration of Rossby waves into the upwelling branch of the diabatic circulation (i.e., the tropical pipe), enhancing mixing. Once the tropical winds are constrained to easterlies in the two models, the models simulate nearly identical mixing.

Even in the constrained SP tests, however, the simulation of mean stratospheric age differs by up to 15%, a difference that remains fairly constant as the resolution is increased in both the horizontal and vertical. The TLP analysis suggests that differences in net diabatic circulation account for the bulk of the age difference (50%, Fig. 9). Differences in mixing still account for a nonnegligible fraction of the difference, but the second largest contribution is the boundary condition at the tropopause, which is primarily due to difference in the overturning circulation in the troposphere.

Our application of the TLP forms a natural extension to the work of Ray et al. (2010), which used the TLP to study the changes in mean circulation and isentropic mixing to study multidecadal ozone variability, explaining the effects of extreme events such as volcanic eruptions on the stratospheric ozone variability. Our framework, which considers full vertical variations of the TLP parameters, provides an effective way to assess the impact of dynamical processes in the stratosphere. It could potentially be used to more accurately diagnose multidecadal changes in the vertical structure of stratospheric circulation and mixing, using satellite-based observations of stratospheric tracers or comprehensive climate models.

Tracing the difference in the diabatic circulation back to the numerical formulation of the dynamical core is beyond the scope of our analysis. The strength of the diabatic circulation is set by the efficiency of large-scale Rossby waves in mixing potential vorticity in the stratosphere, driving the extratropical pump, which lifts mass up into the tropics and back down in the extratropics (Haynes et al. 1991; Holton et al. 1995). The stronger diabatic circulation in GFDL-FV3 is consistent with its slightly weaker climatological polar vortex, in that both are responses to stronger drag on the mean flow by eddies. These differences must relate to differences in the efficiency of wave breaking. The potential for a positive feedback between a weaker vortex, which allows further Rossby wave propagation, which in turn further weakens the vortex, could exacerbate subtle differences in the numerics.

Enhanced wave driving implies more mixing of potential vorticity by wave breaking; assuming a fixed mixing efficiency, this would lead to enhanced mixing of tracer as well, which would partially offset the reduction of aging by the stronger diabatic circulation. In the lower stratosphere (near 390 K), however, the SE core exhibits greater mixing efficiency (Fig. 6b), leading to additional aging of the flow up to about 600 K (Fig. 9). Above this height, the mixing efficiency is comparable in both cores (implying more absolute mixing in the FV3 core), which does partially offset the age difference associated with the weaker diabatic circulation of SE.

Our results confirm the importance of the tropical stratosphere on stratospheric transport (Punge et al. 2009) and the sensitivity of tropical stratospheric climatology to model numerics (Yao and Jablonowski 2015). The momentum balance of the tropical stratosphere is a stiff test for model numerics and is vital for both dynamics (e.g., the quasi-biennial

oscillation) and transport, particularly for the relative isolation of the tropical upwelling region. As minor differences in the tropical winds can have consequences for global stratospheric transport, it is tempting to hope that, given reasonably high resolution, tuning a model to the correct climatological circulation is sufficient to accurately model transport. The manner in which the model state is nudged, however, can have significant impact on the results.

Using the specified dynamics version of Whole Atmosphere Community Climate Model (SD-WACCM), Davis et al. (2020) show that nudging toward the anomalies of the climatological zonal wind and temperature, as opposed to the full zonal winds and temperature profiles, allowed them to capture tropical upwelling trends more accurately in the lower stratosphere. The improvement can, in part, be due to a more accurate representation of wave propagation and hence, potential vorticity trends. Their nudging technique contrast with the uniform (but relatively gentle) nudging of the total tropical wind in this study. While Davis et al. (2020) only used WACCM in their analysis, it is possible that models with different dynamical cores may respond differently to identically imposed nudging, either due to differences in gridscale diffusion or due to differences in wave representation, potentially limiting our interpretation of the remaining transport differences.

Furthermore, Chrysanthou et al. (2019) suggest that the manner in which you correct model biases matters a lot to transport. They found significant differences in diabatic circulation strength among an ensemble of models nudged toward the same reanalysis. Furthermore, M. Linz (2019, personal communication) found a degradation of transport in SD-WACCM relative to free-running WACCM. In both of these cases, there is a danger that nudging a model to the correct climate state leads to an imbalance between model's transport and dynamics. The key is to get the right climatological state *self-consistent* with the underlying dynamics of the model.

Acknowledgments. We acknowledge support of the U.S. National Science Foundation through Grant AGS-1852727 to New York University. We also thank three anonymous reviewers, Marianna Linz, and Olivier Pauluis for constructive criticism on earlier versions of the manuscript.

REFERENCES

Baldwin, M. P., and Coauthors, 2001: The quasi-biennial oscillation. *Rev. Geophys.*, **39**, 179–229, https://doi.org/10.1029/1999RG000073.

Birner, T., and H. Bönisch, 2011: Residual circulation trajectories and transit times into the extratropical lowermost stratosphere. *Atmos. Chem. Phys.*, **11**, 817–827, https://doi.org/10.5194/acp-11-817-2011.

Butchart, N., 2014: The Brewer-Dobson circulation. *Rev. Geophys.*, **52**, 157–184, https://doi.org/10.1002/2013RG000448.

——, and Coauthors, 2018: Overview of experiment design and comparison of models participating in phase 1 of the SPARC Quasi-Biennial Oscillation initiative (QBOi). *Geosci. Model Dev.*, **11**, 1009–1032, https://doi.org/10.5194/gmd-11-1009-2018.

Chrysanthou, A., and Coauthors, 2019: The effect of atmospheric nudging on the stratospheric residual circulation in chemistry–climate models. *Atmos. Chem. Phys.*, **19**, 11 559–11 586, https://doi.org/10.5194/acp-19-11559-2019.

Collins, W. J., and Coauthors, 2017: AerChemMIP: Quantifying the effects of chemistry and aerosols in CMIP6. *Geosci. Model Dev.*, **10**, 585–607, https://doi.org/10.5194/gmd-10-585-2017.

Davis, N. A., S. M. Davis, R. W. Portmann, E. Ray, K. H. Rosenlof, and P. Yu, 2020: A comprehensive assessment of tropical stratospheric upwelling in the specified dynamics Community Earth System Model 1.2.2–Whole Atmosphere Community Climate Model (CESM (WACCM)). *Geosci. Model Dev.*, **13**, 717–734, https://doi.org/10.5194/gmd-13-717-2020.

Dewan, E. M., 1981: Turbulent vertical transport due to thin intermittent mixing layers in the stratosphere and other stable fluids. *Science*, **211**, 1041–1042, https://doi.org/10.1126/science.211.4486.1041.

Donner, L. J., and Coauthors, 2011: The dynamical core, physical parameterizations, and basic simulation characteristics of the atmospheric component AM3 of the GFDL global coupled model CM3. *J. Climate*, **24**, 3484–3519, https://doi.org/10.1175/2011JCLI3955.1.

Engel, A., and Coauthors, 2009: Age of stratospheric air unchanged within uncertainties over the past 30 years. *Nat. Geosci.*, **2**, 28–31, https://doi.org/10.1038/ngeo388.

Eyring, V., and Coauthors, 2007: Multimodel projections of stratospheric ozone in the 21st century. *J. Geophys. Res.*, **112**, D16303, https://doi.org/10.1029/2006JD008332.

Friedl, L., 2007: Aerosols, air quality, and international policy. *36th Int. Seminar on Nuclear War and Planetary Emergencies*, Erice, Italy, Ettore Majorana Foundation, 151–162, https://doi.org/10.1142/9789812709233_0023.

Garny, H., T. Birner, H. Bönisch, and F. Bunzel, 2014: The effects of mixing on age of air. *J. Geophys. Res. Atmos.*, **119**, 7015–7034, https://doi.org/10.1002/2013JD021417.

Gupta, A., E. P. Gerber, and P. H. Lauritzen, 2020: Numerical impacts on tracer transport: A proposed intercomparison test of atmospheric general circulation models. *Quart. J. Roy. Meteor. Soc.*, **146**, 3937–3964, https://doi.org/10.1002/qj.3881.

Hall, T. M., and R. A. Plumb, 1994: Age as a diagnostic of stratospheric transport. *J. Geophys. Res.*, **99**, 1059–1070, https://doi.org/10.1029/93JD03192.

——, D. W. Waugh, K. A. Boering, and R. A. Plumb, 1999: Evaluation of transport in stratospheric models. *J. Geophys. Res.*, **104**, 18 815–18 839, https://doi.org/10.1029/1999JD900226.

Haynes, P., and E. Shuckburgh, 2000: Effective diffusivity as a diagnostic of atmospheric transport: 1. Stratosphere. *J. Geophys. Res.*, **105**, 22 777–22 794, https://doi.org/10.1029/2000JD900093.

——, M. E. McIntyre, T. G. Shepherd, C. J. Marks, and K. P. Shine, 1991: On the ''downward control'' of extratropical diabatic circulations by eddy-induced mean zonal forces. *J. Atmos. Sci.*, **48**, 651–678, https://doi.org/10.1175/1520-0469(1991)048<0651:OTCOED>2.0.CO;2.

Held, I. M., and M. J. Suarez, 1994: A proposal for the intercomparison of the dynamical cores of atmospheric general circulation models. *Bull. Amer. Meteor. Soc.*, **75**, 1825–1830, https://doi.org/10.1175/1520-0477(1994)075<1825:APFTIO>2.0.CO;2.

Holton, J. R., P. H. Haynes, M. E. McIntyre, A. R. Douglass, R. B. Rood, and L. Pfister, 1995: Stratosphere-troposphere exchange. *Rev. Geophys.*, **33**, 403–439, https://doi.org/10.1029/95RG02097.

Kida, H., 1983: General circulation of air parcels and transport characteristics derived from a hemispheric GCM: Part 2. Very long-term motions of air parcels in the troposphere and stratosphere. *J. Meteor. Soc. Japan*, **61**, 510–523, https://doi.org/10.2151/jmsj1965.61.4_510.

Kolonjari, F., and Coauthors, 2018: Assessing stratospheric transport in the CMAM30 simulations using ACE-FTS measurements. *Atmos. Chem. Phys.*, **18**, 6801–6828, https://doi.org/10.5194/acp-18-6801-2018.

Lauritzen, P. H., and Coauthors, 2018: NCAR release of CAM-SE in CESM2.0: A reformulation of the spectral element dynamical core in dry-mass vertical coordinates with comprehensive treatment of condensates and energy. *J. Adv. Model. Earth Syst.*, **10**, 1537–1570, https://doi.org/10.1029/2017MS001257.

Lindzen, R. S., and M. Fox-Rabinovitz, 1989: Consistent vertical and horizontal resolution. *Mon. Wea. Rev.*, **117**, 2575–2583, https://doi.org/10.1175/1520-0493(1989)117<2575:CVAHR>2.0.CO;2.

Linz, M., R. A. Plumb, E. P. Gerber, and A. Sheshadri, 2016: The relationship between age of air and the diabatic circulation of the stratosphere. *J. Atmos. Sci.*, **73**, 4507–4518, https://doi.org/10.1175/JAS-D-16-0125.1.

——, ——, ——, F. J. Haenel, G. Stiller, D. E. Kinnison, A. Ming, and J. L. Neu, 2017: The strength of the meridional overturning circulation of the stratosphere. *Nat. Geosci.*, **10**, 663–667, https://doi.org/10.1038/ngeo3013.

——, ——, A. Gupta, and E. P. Gerber, 2021: Stratospheric adiabatic mixing rates derived from the vertical gradient of age of air. *J. Geophys. Res. Atmos.*, https://doi.org/10.1029/2021JD035199, in press.

McIntyre, M. E., and T. N. Palmer, 1983: Breaking planetary waves in the stratosphere. *Nature*, **305**, 593–600, https://doi.org/10.1038/305593a0.

Molina, M. J., and F. S. Rowland, 1974: Stratospheric sink for chlorofluoromethanes: Chlorine atom-catalysed destruction of ozone. *Nature*, **249**, 810–812, https://doi.org/10.1038/249810a0.

Nakamura, N., 1996: Two-dimensional mixing, edge formation, and permeability diagnosed in an area coordinate. *J. Atmos. Sci.*, **53**, 1524–1537, https://doi.org/10.1175/1520-0469(1996)053<1524:TDMEFA>2.0.CO;2.

Neu, J. L., and R. A. Plumb, 1999: Age of air in a "leaky pipe" model of stratospheric transport. *J. Geophys. Res.*, **104**, 19 243–19 255, https://doi.org/10.1029/1999JD900251.

Orbe, C., M. Holzer, L. M. Polvani, and D. Waugh, 2013: Air-mass origin as a diagnostic of tropospheric transport. *J. Geophys. Res. Atmos.*, **118**, 1459–1470, https://doi.org/10.1002/jgrd.50133.

Pauluis, O., A. Czaja, and R. Korty, 2009: The global atmospheric circulation in moist isentropic coordinates. *J. Climate*, **23**, 3077–3093, https://doi.org/10.1175/2009JCLI2789.1.

Plumb, R. A., 1996: A "tropical pipe" model of stratospheric transport. *J. Geophys. Res.*, **101**, 3957–3972, https://doi.org/10.1029/95JD03002.

——, 2002: Stratospheric transport. *J. Meteor. Soc. Japan*, **80**, 793–809, https://doi.org/10.2151/jmsj.80.793.

——, 2007: Tracer interrelationships in the stratosphere. *Rev. Geophys.*, **45**, RG4005, https://doi.org/10.1029/2005RG000179.

——, and R. C. Bell, 1982: A model of the quasi-biennial oscillation on an equatorial beta-plane. *Quart. J. Roy. Meteor. Soc.*, **108**, 335–352, https://doi.org/10.1002/qj.49710845604.

——, and M. K. W. Ko, 1992: Interrelationships between mixing ratios of long-lived stratospheric constituents. *J. Geophys. Res.*, **97**, 10 145–10 156, https://doi.org/10.1029/92JD00450.

Polvani, L. M., and P. J. Kushner, 2002: Tropospheric response to stratospheric perturbations in a relatively simple general circulation model. *Geophys. Res. Lett.*, **29**, 1114, https://doi.org/10.1029/2001GL014284.

Punge, H. J., P. Konopka, M. A. Giorgetta, and R. Müller, 2009: Effects of the quasi-biennial oscillation on low-latitude transport in the stratosphere derived from trajectory calculations. *J. Geophys. Res.*, **114**, D03102, https://doi.org/10.1029/2008JD010518.

Ray, E. A., and Coauthors, 2010: Evidence for changes in stratospheric transport and mixing over the past three decades based on multiple data sets and tropical leaky pipe analysis. *J. Geophys. Res.*, **115**, D21304, https://doi.org/10.1029/2010JD014206.

——, F. L. Moore, K. H. Rosenlof, D. A. Plummer, F. Kolonjari, and K. A. Walker, 2016: An idealized stratospheric model useful for understanding differences between long-lived trace gas measurements and global chemistry-climate model output. *J. Geophys. Res. Atmos.*, **121**, 5356–5367, https://doi.org/10.1002/2015JD024447.

Sparling, L. C., J. A. Kettleborough, P. H. Haynes, M. E. McIntyre, J. E. Rosenfield, M. R. Schoeberl, and P. A. Newman, 1997: Diabatic cross-isentropic dispersion in the lower stratosphere. *J. Geophys. Res.*, **102**, 25 817–25 829, https://doi.org/10.1029/97JD01968.

Townsend, R. D., and D. R. Johnson, 1985: A diagnostic study of the isentropic zonally averaged mass circulation during the first GARP global experiment. *J. Atmos. Sci.*, **42**, 1565–1579, https://doi.org/10.1175/1520-0469(1985)042<1565:ADSOTI>2.0.CO;2.

Waugh, D., and T. Hall, 2002: Age of stratospheric air: Theory, observations, and models. *Rev. Geophys.*, **40**, 1010, https://doi.org/10.1029/2000RG000101.

——, and Coauthors, 1997: Mixing of polar vortex air into middle latitudes as revealed by tracer-tracer scatterplots. *J. Geophys. Res.*, **102**, 13 119–13 134, https://doi.org/10.1029/96JD03715.

Yamada, R., and O. Pauluis, 2015: Momentum balance and Eliassen–Palm flux on moist isentropic surfaces. *J. Atmos. Sci.*, **73**, 1293–1314, https://doi.org/10.1175/JAS-D-15-0229.1.

Yao, W., and C. Jablonowski, 2015: Idealized quasi-biennial oscillations in an ensemble of dry GCM dynamical cores. *J. Atmos. Sci.*, **72**, 2201–2226, https://doi.org/10.1175/JAS-D-14-0236.1.