GISS Model E2.2: A Climate Model Optimized for the Middle Atmosphere. Part 2: Validation of Large-Scale Transport and Evaluation of Climate Response

Clara Orbe[1], David Rind[1], Jeffrey Jonas[2,1], Larissa Nazarenko[2,1], Greg Faluvegi[2,1], Lee T. Murray[3], Drew T. Shindell[4], Kostas Tsigaridis[2,1], Tiehan Zhou[2,1], Maxwell Kelley[5,1], Gavin A. Schmidt[1]

[1]NASA Goddard Institute for Space Studies, 2880 Broadway New York, NY 10025
[2]Center for Climate Systems Research, Earth Institute, Columbia University, New York NY
[3]Department of Earth and Environmental Sciences, University of Rochester, Rochester NY
[4]Nicholas School of the Environment, Duke University, Durham NC
[5]SciSpace LLC, New York NY

Key Points:

- The stratospheric transport circulation is evaluated for the new GISS "high-top" CMIP6 climate model.
- Stratospheric mean ages are significantly improved compared to the lower vertical resolution version of ModelE.
- The stratospheric transport response to increased CO_2 is approximately linear and correlated with the magnitude of surface warming.

Corresponding author: Clara Orbe, clara.orbe@nasa.gov

Abstract

Here we examine the large-scale transport characteristics of the new "Middle Atmosphere" NASA Goddard Institute for Space Studies climate model (E2.2). First we evaluate the stratospheric transport circulation in historical atmosphere-only simulations integrated with interactive trace gas and aerosol chemistry. Compared to lower vertical resolution model versions, E2.2 exhibits improved tropical ascent and older stratospheric mean ages that are more consistent with observed values. In the troposphere, poleward transport to the Arctic and interhemispheric mean ages in E2.2 are comparable to models participating in the Chemistry Climate Modeling Initiative.

In addition to validating E2.2 we also assess its "transport sensitivity" using the coupled atmosphere-ocean abrupt 4xCO_2 and transient 1%CO_2 simulations submitted to the Coupled Model Intercomparison Project, Phase 6, along with a 2xCO_2 simulation used to evaluate the linearity of the transport circulation's response to increased CO_2. We show that decreases (increases) in a stratospheric mean age (idealized surface loss) tracer scale linearly with increased lower stratospheric upwelling, which also increases linearly with warming tropical sea surface temperatures (SSTs). Abrupt 2xCO_2 and 4xCO_2 experiments constrained with (fixed) pre-industrial SSTs are also used to quantify the relative importance of rapid adjustments versus SST feedbacks to the transport circulation responses in the model. Finally, sensitivity experiments are presented to illustrate the impact of changes in the convective parameterization on stratospheric transport.

1 Introduction

It is now well appreciated that the stratosphere plays an important role in shaping the chemical, dynamical and radiative properties at Earth's surface on a range of timescales. On climatic timescales these include the modulation of the Southern Hemisphere midlatitude jet and Southern Ocean ventilation changes by stratospheric ozone depletion (e.g., Polvani et al. (2011); Waugh et al. (2013)). On seasonal timescales, the stratosphere also influences surface weather over both the extra-tropics (e.g. Gerber et al. (2012); Scaife et al. (2016); Seviour et al. (2014)) and in the tropics (Yoo & Son, 2016). While some of these influences are almost exclusively dynamical in nature, others are largely mediated by atmospheric composition, most notably through changes in stratospheric ozone, which influences lower stratospheric temperature gradients and, via thermal wind balance, the tropospheric midlatitude jet streams. Therefore, in order to capture the full effect of the stratosphere on surface climate trends and variability it is important that models properly simulate stratospheric composition including, but not limited to, ozone, water vapor, and stratospheric aerosols.

The global-scale characteristics of stratospheric tracers (and their mutual relationships) are dominated not only by mean diabatic advection (upwelling in the tropics, downwelling in the surf zone and the vortex) but also by rapid isentropic stirring within the surf zone (Plumb, 2002). Therefore, in order to properly simulate stratospheric composition it is important that models not only accurately represent dynamical measures like the residual mean (or Transformed Eulerian Mean) circulation (D. G. Andrews et al., 1987) but, also, the integrated effects of advection and isentropic mixing, as captured through "tracer-independent" measures of the transport circulation (Holzer & Hall, 2000) like the stratospheric mean age (Hall & Plumb, 1994).

In Part I of this study Rind et al. (2020) (hereafter R20) presented a comprehensive overview of the new NASA Goddard Institute for Space Studies (GISS) Middle Atmosphere (MA) Model E2.2. While that study examined a range of key dynamical features of E2.2, both with respect to its mean state (e.g. Hadley Cell strength, zonal winds) as well as its variability (e.g. Madden-Julian Oscillation, the Quasi-Biennial Oscillation (QBO), Stratospheric Sudden Warmings), here we present an overview of the large-scale transport characteristics of the model. As in that study we focus primarily on the stratosphere, which differs most compared to the lower vertical resolution version of GISS ModelE (E2.1) (Kelley et al., 2019) with which we compare directly in addition to comparing against observations. In particular, the stratospheric transport circulation in E2.2 is expected to reflect not only the large-scale (largely advective) dynamical features discussed in Part I, but also the representation of isentropic mixing within the middle and lower stratosphere, which is sensitive to vertical resolution as it depends on how well meridional and vertical tracer gradients are represented.

For sake of brevity we focus our evaluation of E2.2 around the atmosphere-only (AMIP) historical and coupled atmosphere-ocean increased carbon dioxide (CO_2) simulations that have been submitted as part of the DECK to the Coupled Model Intercomparison Project, Phase 6 (CMIP6) (Eyring et al., 2016), using the former to validate the model's large-scale transport characteristics and the latter to evaluate the climate response of the transport circulation. Although not requested for any particular "MIP", all simulations presented here were integrated carrying a range of idealized tracers that provide canonical measures for evaluating large-scale transport as used in previous (non-CMIP) intercomparisons of chemistry climate models, including the Stratospheric Processes and their Role in Climate (SPARC) CCMVal (CCMVal, 2010) and, more recently, the Chemistry-Climate Model Initiative (CCMI) (Eyring et al., 2013).

Our discussion of the climate change response in the stratospheric and tropospheric transport circulations is in the context of the abrupt CO_2 forcing simulations which, despite their simplicity, afford a mechanistic look into the response characteristics of models that can be unambiguously attributed to an increase in CO_2 concentrations. This simple forcing lens is especially important for understanding the large-scale transport response to climate change, which has been relatively unexplored in previous studies that have either focused almost exclusively on "dynamical sensitivity" (Grise & Polvani, 2016; Chemke & Polvani, 2019; Menzel et al., 2019) or on the large-scale transport circulation response to changes in both CO_2 (and other greenhouse gases (GHGs)) and ozone depleting substances (ODS) (Doherty et al., 2017; Abalos et al., 2019). The CO_2-induced response of the large-scale transport circulation therefore remains poorly understood.

We begin by comparing in Section 3 various chemical and transport measures in the CMIP6 historical simulations of E2.2 over the recent satellite period with observations, when available. In those cases where comparisons with observations are not possible we compare E2.2 directly with results from the CCMI models presented in Orbe et al. (2019) in an effort to place Model E2.2 in the broader context of similar high vertical resolution chemistry climate models (CCMs). Then we present different measures of "transport sensitivity" in E2.2 in Section 4 along with a discussion of the sensitivity of different aspects of the stratospheric transport circulation to changes in (parameterized) convection and how this informed model development (Section 5). Conclusions are then presented in Section 6.

2 Analysis Approach

Here we review the models, simulations and transport and chemical measures used in our analysis. The observational products against which the simulations are compared are also discussed.

2.1 Models: E2.2(-AP) and E2.1

Our analysis focuses on the NASA GISS "Middle Atmosphere" model E2.2, which was documented extensively in R20 in terms of its key radiative and dynamical properties. Unlike previous MA versions of ModelE, E2.2 has contributed to CMIP6, with all DECK simulations having thus far been submitted to the Earth System Grid Federation (`https://esgf.llnl.gov`), for the case of the coupled atmosphere-ocean model utilizing non-interactive (NINT) chemistry. Interactive versions of the E2.2 DECK are currently processing and will also be made available upon their completion.

As described in R20, E2.2 consists of 102 vertical levels spanning the surface up to 0.002 hPa, as compared to the lower vertical resolution of ModelE (E2.1), which consists of 40 levels extending up to 0.1 hPa. Orographic and non-orographic gravity wave drag (GWD) is parameterized following Lindzen (1984) and Rind et al. (1988), producing in E2.2 a Quasi-Biennial Oscillation that compares well with observations as well as improved stratospheric polar vortex variability (R20, Ayarzagüena et al. (2020)). We refer the reader to R20 for an in-depth discussion of the model.

Among the different model versions discussed in R20 here we focus on the "Altered-Physics" (-AP) version E2.2-AP because this is the configuration that was submitted to CMIP6 and presented in Ayarzagüena et al. (2020). While this version does differ from the "standard" model version E2.2 in certain respects (i.e. convective mass flux profiles, high cloud cover, planetary albedo, shortwave absorbed at the surface) the climatologies of both model versions agree overall quite well, especially with respect to their stratospheric transport properties, as discussed in Sections 3 and 4.

Finally, R20 reviewed aspects of the dynamical parameterizations that differ between E2.1 and E2.2, including, among others, incorporation–and subsequent tuning– of an "efficiency factor" relating convection to parameterized gravity wave momentum fluxes (see R20 for more). In addition, while the gas phase chemistry formulation in E2.2 is that used in E2.1, changes in the dynamics and thermodynamic structure associated with the higher model top in E2.2 necessitated a re-tuning of some aspects of the photolysis code. In particular, in E2.1 overhead ozone above the top of the chemistry is taken to be constant, whereas in E2.2 it is given a spatial variation to match that of ozone at the top layer of chemistry. In addition, certain photolysis rates tunings at short wavelengths ($<$ 200 nm) for N_2O and O_2, which in E2.1 corrected for stratospheric circulation-induced biases in high latitude NO_x and O_3, were disabled in E2.2.

2.2 Experiments

2.2.1 E2.2-AP CMIP6 DECK Simulations

The bulk of our analysis presents results from the CMIP6 DECK experiments performed using E2.2-AP. For purposes of validating the model we begin by first discussing the results from the AMIP historical simulations constrained with observed sea surface temperatures (SSTs) and prescribed sea ice concentrations (SICs) (Table 1, rows 1-2). Specifically, we utilize the "OMA" (AMIP) version of E2.2-AP described in R20 in which aerosols and trace gases are calculated interactively using the OMA ("one-moment aerosol") scheme ("TCADI" in CMIP5). Results from three members of the E2.2-AP ensemble are presented (row 1). In addition, one member of a 5-member E2.2 OMA AMIP ensemble is also presented (row 2) in order to show that the "Altered Physics" version of the model exhibits quite similar stratospheric transport characteristics to those of the "standard" E2.2 version. Note that we only use one member for the latter since the transport differences among the three E2.2-AP members is found to be quite small, as demonstrated in Section 3.

After using the AMIP OMA E2.2-AP simulations for validation against the observations, we then present an analysis of the "transport sensitivity" of the "Coupled" model from R20, using results from the E2.2-AP DECK coupled atmosphere-ocean non-interactive pre-industrial control and abrupt and transient CO_2 simulations (Table 1, rows 4-7). For each simulation only one ensemble member was run for CMIP6 (and shown herein). In particular, we examine the Pre-Industrial (PI) control and "branching" abrupt $2xCO_2$ and $4xCO_2$ and transient $1\%CO_2$ experiments. In this study we focus only on runs coupled to the GISS Ocean v1 (GO1) (i.e. "-G" in CMIP6 notation, "-R" in CMIP5). Although E2.2-AP DECK simulations were also performed using the HYCOM dynamical ocean (i.e. E2.2-H) they did not integrate the passive tracers that underlie the bulk of our analysis and so only E2.2-AP-G results are presented here. Finally, abrupt 2xCO_2 and 4xCO_2 experiments constrained with (fixed) pre-industrial SSTs (FIXSST) are also used to quantify the relative importance of rapid adjustments versus SST feedbacks (Table 1, rows 8-9).

2.2.2 E2.1 CMIP6 Historical Simulation

In addition to validating E2.2 against observations we also compare with the results from two members of the E2.1 historical AMIP ensemble presented in R20 that was also submitted to CMIP6 (Table 1, row 3). That simulation, which uses the same OMA chemical mechanism, is identical to that presented in Kelley et al. (2019) with the exception that additional idealized tracer diagnostics were integrated in order to facilitate transport comparisons with E2.2. Note that these tracers, as described in Section 2c, are passive and do not interact with the model's physics and/or dynamics.

2.2.3 E2.2 Sensitivity Experiments

Unlike some earlier iterations of the MA version of ModelE (i.e. Model III discussed in Rind et al. (2014)), E2.2 employs the same underlying physics as in E2.1. There are important departures, however. In particular, while both versions use the same deep convection parameterization (Del Genio et al., 2007), with updates in both models designed to enhance the simulation of the MJO (Kim et al., 2013; Kelley et al., 2019), some additional modifications related to how condensate evaporation in convective updrafts interacts with the wider, non-convecting part of the GCM grid box were incorporated in E2.2 and not in E2.1 (R20). The implications of that parameter choice were discussed in R20 and will not be presented here as they do not bear immediate relevance to stratospheric transport.

By comparison, the other main set of changes to the convective parameterization – those that distinguish the "standard" versus "AP" versions of E2.2 – will be briefly presented in Section 5 as the experiments that were conducted in the process of developing E2.2-AP illustrate how convection changes can affect the stratospheric transport characteristics of the model. In deciding on the final version of E2.2-AP we carried out several sensitivity experiments utilizing both AMIP and coupled atmosphere-ocean configurations, only a subset of which is presented here (Table 1, rows 10-11). Our motivation for presenting these results is not to make any direct inferences about specific aspects of the convective parameterization (which likely will be model dependent) but, rather, to illustrate interesting relationships between convection and stratospheric transport that we observed in the process of developing E2.2-AP that might be more generally applicable to other models.

2.2.4 Chemistry Climate Modeling Initiative (CCMI) Simulations

Finally, in order to place the E2.2 OMA (AMIP) results in the broader context of similar high-top stratosphere-resolving models we compare with the transport evaluations for the CCMI models presesented in Orbe et al. (2019). Specifically, we compare against the results from the REF-C1 free-running experiments spanning 1979-2010 which were constrained with observed SSTs and SICs and which integrated a broad set of idealized passive tracers, later implemented in E2.2 as described next.

2.3 Chemical and Transport Measures

To evaluate stratospheric transport in E2.2 we use a combination of both real (i.e. methane (CH_4), nitrous oxide (N_2O), water vapor (H_2O) and ozone (O_3)) and idealized tracers, including the stratospheric mean age of air (Γ_{STRAT}) and an annually periodic oscillating tracer (χ_{tape}) (Hall et al., 1999; Orbe et al., 2017) (Table 2). The latter is used to evaluate tropical ascent and is implemented by prescribing a sinusoid in mixing ratio over 10°S-10°N at 100 mb that has a maximum during October, consistent with the seasonality of water vapor-based estimates of the tape recorder at the tropical tropopause (Mote et al., 1996) and consistent with the implementation in Orbe et al. (2017). At the same time, the mean age provides an integrated measure of the time since air was last at the tropical tropopause. Though not "tracer-independent" (Holzer & Hall, 2000), as is the case for Γ_{STRAT}, the gradients of CH_4 and N_2O also provide information about the relative importance of ascent versus in-mixing within the tropical pipe and are common measures for assessing stratospheric transport in models (Eyring et al. (2006)).

In addition to the stratospheric mean age we also investigate transport from the stratosphere to the troposphere through analysis of the residence time of (simulated) bomb-produced ^{14}C (Prather & Remsberg, 1993; Rind et al., 1999). The initial conditions for the release are taken from Johnston (1989) for the year October 1963, with input peaking at around 20 km at northern mid-latitudes. The lower boundary is varied as in the following prescription (see also Prather and Remsberg (1993)):

$$\chi_{14\mathrm{C}}(\mathbf{r},t) = \begin{cases} 73.0 - 0.27823t - 3.45648e^{-3t^2} + 4.21159e^{-5t^3} & \text{over the NH} \\ 44.5 + 1.02535t - 2.13565e^{-2t^2} + 8.61853e^{-5t^3} & \text{over the SH} \end{cases}$$

where t refers to the number of months after October 1963, the units are 10^5 molecules ^{14}C g-1 of air and NH and SH refer to the Northern and Southern hemispheres, respectively.

Even for cases when the stratospheric mean age and mean residence time are defined with respect to the same stratospheric entry boundary condition (i.e. the tropical tropopause), the two timescales capture physically distinct aspects of stratospheric transport, with the mean age referring to the population of fluid elements that exits the stratosphere, whereas the residence time characterizes the fluid elements that reside in the stratosphere (Holzer et al., 2012). Hall and Waugh (2000) found that they are correlated, but imperfectly so, particularly for the case of release sources occurring in the mid-latitude lower stratosphere, for which variations in the mid-latitude tropopause height strongly affect the residence time but not the mean age.

After analyzing stratospheric transport we then use a combination of other idealized tracers to evaluate different aspects of large-scale tropospheric transport. In order to place E2.2 in the broader context of other CCMs we use tracers that were also implemented in the CCMI REF-C1 integrations and evaluated in Orbe et al. (2019). In particular, we examine two idealized loss tracers that are emitted over the Northern Hemisphere (NH) midlatitude surface (30°N-50°N), $\chi_{\mathrm{NH},5}$ and $\chi_{\mathrm{NH},50}$, and decay uniformly with loss frequencies of 5 days^{-1} and 50 days^{-1}, respectively.

In addition, we consider a tropospheric mean age tracer Γ_{NHMID} (not to be confused with Γ_{STRAT}) that describes the mean time since air was at the NH midlatitude surface. Unlike other measures of interhemispheric transport, the mean age can be calculated for locations throughout the troposphere and thus provides more information on transport times than the interhemispheric exchange time (Geller et al., 1997; Levin & Hesshaimer, 1996), which only quantifies the transport between hemispheres (Waugh et al., 2013). In order to facilitate comparisons of simulated values of Γ_{NHMID} with observations we also integrate a sulfur hexafluoride (SF_6) tracer using the same power grid source distribution implemented in Rind et al. (1999) and subject to increases of 0.3 pptv/year.

From SF_6 we then define an SF_6 age (a_{SF6}) as the time lag satisfying

$$\chi(\mathbf{r}, t) = \chi_0(t - a_{SF6}) \quad (1)$$

where $\chi(\mathbf{r}, t)$ refers to the SF_6 concentration at a location $\mathbf{r}$ and field time t and χ_0 refers to the concentration of SF_6 over the source region (here, the NH midlatitude surface). Note that, unlike for the case of the Γ_{NHMID} tracer, which was implemented identically as in the CCMI models, the emissions for the SF_6 tracer in ModelE are different than the more recent emissions distributions used in CCMI, which were taken from EDGAR v4.2 (Yang et al., 2019). Therefore, while the tropospheric mean age tracer (Γ_{NHMID}) is used to compare E2.2(-AP) and E2.1 with the CCMI models, the SF_6 age tracer is used for comparing against observations.

While we use Γ_{NHMID} and a_{SF6} as integrated measures of interhemispheric transport, we also examine subtropical convection through use of a radon tracer (Rn-222), the vertical distribution of which is governed primarily by fast local mixing processes, owing to its surface source derived from the decay of uranium ore (U-238) in soils and its half-life of 3.8 days. Unfortunately, very few observed vertical profiles exist in the tropics or subtropics for comparisons with models. Therefore, our evaluations of simulated radon in the subtropics are limited to profiles at only one location (Rani et al., 2014), as discussed in Section 3. Additional comparisons of Rn-222 with the observed profiles over northern midlatitudes from Murray et al. (2014) are also performed.

Finally, we consider the distribution of the idealized tracer "e90" to quantify transport in the upper troposphere/lower stratosphere (UTLS), with regards to both its interannual variability (Abalos et al., 2017) as well as its response to climate change (Abalos et al., 2019). Specifically, e90 is emitted at the Earth's surface and decays uniformly with a lifetime of 90 days^{-1} throughout the atmosphere.

2.4 Observational Products

Various observational products are used to evaluate the chemical and transport characteristics of the E2.2 historical simulations (Table 2, col. 3). Several of these were used in CCMVal (2010), and are included here to ensure consistency in our (stratospheric) transport evaluation with that of other models, except in cases where new observations correct for known biases in the older products.

In particular, the simulated fields of CH_4, O_3 and H_2O are compared with the climatologies derived for 1991-2002 from the Halogen Occultation Experiment (HALOE) on board the Upper Atmosphere Research Satellite (UARS) (Grooß & Russell, 2005). Comparisons of simulated N_2O are made against 2005-2015 climatologies derived from the Microwave Limb Sounder (MLS) on the Earth Observing System (EOS) Aura satellite. We use the 190-GHz retrieval from version 4.2 because the 640-GHz data set ends in summer 2013 due to the failure of the N_2O primary band. Note that any recent lower stratospheric changes in N_2O (those occurring after 2015) are not considered in this study (Personal Communication with Krzysztof Wargan). For ozone, in addition to the stratospheric profiles from HALOE, total column ozone fields are also evaluated against the Total Ozone Mapping Spectrometer (TOMS) and the Ozone Monitoring Instrument (OMI) (TOMS for years 2000-2004 and OMI for 2005-2010) (R. McPeters et al., 2008). Historical trends in simulated total column ozone (TCO) are compared against the ground-based measurements based on the Dobson and Brewer spectrophotometer and filter ozonometer observations available from the World Ozone and UV Data Centre (WOUDC) updated from (Fioletov et al., 2002). In addition to the ground-based observations, which date back to 1964, we also compare simulated TCO values to the monthly mean zonally averaged SBUV (Version 8.6) merged ozone dataset extending back to January 1970 (R. D. McPeters et al., 2013).

In addition to our evaluations of the chemical tracers, simulated values of the stratospheric mean age (Γ_{STRAT}) are compared first against meridional age profiles derived from in-situ aircraft measurements of carbon dioxide (CO_2), averaged in 2.5° latitude bins over the altitude range 19.5-21.5 km (Boering et al. (1996), see also Figure 5 in Hall et al. (1999)). Vertical profiles of simulated Γ_{STRAT} are also compared in the tropics against the average of in situ-based estimates derived separately from CO_2 and SF_6 sampled over 10°S-10°N between 15.2-34.2 km and 15.2-34 km, respectively. Over midlatitudes only the CO_2-based age profiles are used, which apply to latitudes spanning 34°N-44°N and altitudes between 11.1 and 35.1 km (A. E. Andrews et al., 2001; Engel et al., 2009). HALOE H_2O fields are used to obtain the observational-based tape recorder (χ_{tape}) phase lag values presented in this study and mirror those shown in Hall et al. (1999) (See their Figure 16).

Finally, simulated values of the SF_6 age (a_{SF6}) are compared against the observed profiles that were presented in Waugh et al. (2013) (see their Figure 3). In the calculation of a_{SF6} from the observations, χ_0 (from Equation 1) is taken to be the average of SF_6 obtained from surface flask measurements from three NH midlatitude stations (Mace Head (53.5°N), Niwot Ridge (40°N) and THD (Trinidad Head, 41°N)) from the NOAA Halocarbons and other Atmospheric Trace Species (HATS) group. The SF_6 age at southern latitudes is then calculated using a combination of measurements from both ground stations, including HATS as well as the discrete air samples collected from the NOAA Carbon Cycle Greenhouses Gases (CCGG) group, and commercial ships. We refer the reader to Waugh et al. (2013) for more details.

3 Transport Evaluation of E2.2 CMIP6 Historical AMIP Simulations

We begin by presenting climatologies of various stratospheric constituents that can be directly compared with observations and, in turn, provide an indirect measure of how well the transport circulation is being simulated in the model. We then present a comparison of the tracer-independent measures of transport (i.e. both stratospheric and tropospheric mean ages, the tape-recorder phase lag) for which some observational constraints are available and, in cases where they are not, comparisons are made directly with the CCMI models.

3.1 Stratospheric Transport Circulation

3.1.1 Annual Climatological Stratospheric CH_4, N_2O, O_3 and H_2O

Zonal mean comparisons of methane (CH_4) between E2.2-AP and HALOE (Fig. 1a (top), Supplementary Fig. 1a (top)) show good agreement throughout the lowermost stratosphere (30-100 hPa) over both the tropics and extra-tropics, in terms of both mean values in the tropics and in terms of the meridional slopes of tracer isopleths over the subtropics and midlatitudes. (Note that only latitudes between 60°S-60°N are shown for HALOE, owing to uncertainties in the measurements over the poles, which vary with season (see Grooß and Russell (2005) for more). Overall, the absolute values and, in particular, isopleth shapes over the extra-tropics in E2.2-AP are improved compared to in E2.1 (Fig. 1a, bottom), although their gradients are still somewhat weaker compared to the observations. In E2.1 the presence of much weaker gradients in the subtropics and midlatitudes in both hemispheres is indicative of a leakier tropical pipe. This excessive leakiness in E2.1 is consistent with much larger transient eddy kinetic energy biases in that model throughout the subtropical middle and lower stratosphere, as compared to E2.2-AP (not shown).

Stronger meridional gradients in the lower stratosphere in E2.2-AP are also exhibited in other tracers, including N_2O (Fig. 1b, Supplementary Fig. 1b). In addition to improved gradients, the absolute values of N_2O in E2.2-AP also exhibit overall much better agreement with the observations, compared to E2.1. One exception, however, is the tropical middle stratosphere (∼10-20 hPa) where E2.2-AP exhibits a low (∼10%) bias (featured also in methane), that is not seen in E2.1, which exhibits values at 20 hPa (24×10^{-8} V/V) that compare well with MLS. As discussed in the next section, these localized tropical mid-stratospheric biases in E2.2-AP most likely reflect excessive in-mixing between the tropics and subtropics in that region.

Moving next to ozone, we find that O_3 concentrations in both E2.1 and E2.2-AP compare well with observed values over all latitudes and for pressures greater than 30 hPa (Fig. 1c, Supplementary Fig. 1c). In the lowermost stratosphere (i.e. 50-100 hPa) the ozone concentrations in both models are similar, albeit smaller in E2.2-AP poleward of 40°S/N and larger in the tropics and subtropics. The smaller stratospheric ozone values in E2.2-AP over latitudes poleward of 40°S/N most likely contribute to the overall smaller total ozone burdens in that model, compared to E2.1, manifest both in the seasonal cycle and historical trends as discussed in Sections 3.3 and 3.5, respectively.

Above 30 hPa, the ozone values in E2.2-AP are too low in the tropics, a bias also exhibited in E2.2 (not shown). One way to interpret this low bias, which is not exhibited in E2.1, is in terms of greater tropical upwelling in this region in E2.1 compared to E2.2(-AP), which is manifest in stratospheric mean age differences between the models (Fig. 2) as discussed more in the next section. However, while this explains the O_3 differences between E2.2(-AP) and E2.1, it does not explain why E2.2(-AP) is biased low, compared to the observations. Moreover, the fact that this (relatively localized) ozone bias also occurs in a region of low CH_4 and N_2O, indicates that it is also likely not fundamentally driven by photolysis differences between the models but, rather, more likely reflects a more general dynamical circulation bias in E2.2.

To this end, further inspection of the zonal winds in this region (Figures 5 and 35 in R20) reveals a localized region of anomalous westerlies in both E2.2-AP and E2.2 not present in E2.1 during boreal winter. Compared to the overall climatological wind and temperature distributions in the stratosphere, which are much improved in E2.2(-AP) compared to E2.1 (R20), this bias is small in amplitude and regional in nature. Nonetheless, its presence may have an impact on the transport properties in that region. In particular, while their origins are not well understood, in addition to having a direct advective impact on tracer transport in the tropics, these localized wind biases are also associated with changes in meridional potential vorticity (PV) gradients which can directly impact mixing into and out of the tropical pipe. Indeed, as shown in Eichinger et al. (2020), the incorporation of non-orographic gravity wave drag in simulations using the EMAC chemistry climate model has a significant impact on the strength of PV gradients (and associated mixing) in this region (i.e. the tropics spanning pressures between 2-20 hPa) (See their Figure 8). Therefore, the incorporation of additional non-orographic GWD sources in E2.2(-AP) may also impact tropical transport indirectly through changes in mixing, not only through the direct (advective) changes associated with explicit simulation of the QBO, as discussed later.

Finally, comparisons of stratospheric water vapor show good agreement between E2.2-AP and HALOE, albeit with a slightly low bias (∼5%) over the northern midlatitude stratosphere and a wet bias (∼5%) over the tropical lower stratosphere (Fig. 1d, Supplementary Fig. 1d). The water vapor fields in E2.2-AP represent an improvement over E2.1 in terms of both absolute magnitudes (E2.1 is biased low by ∼10-20%), as well as in terms of meridional gradients over the extra-tropical stratosphere. The larger values in E2.2-AP in the tropical lower stratosphere are most likely associated with a warmer tropical tropopause cold point temperature, which is warmer by ∼1-2 degrees in E2.2-AP, compared to E2.1, which is biased cold (see Figure 4 in R20 which also shows that E2.2-AP is biased ∼0.5-1.0 degrees warm relative to reanalysis fields). Furthermore, we note that E2.2, which is still warmer in the tropopause region, exhibits a slightly wetter bias (not shown). Thus, while details of the processes that control stratospheric water vapor remain controversial (Danielsen, 1993; Sherwood & Dessler, 2000; Holton & Gettelman, 2001), the lower tropical stratospheric biases in the different versions of ModelE do seem to be very much tethered to their respective climatological cold point temperatures, consistent with previous studies (Randel et al., 2004).

3.1.2 Annual Climatological Mean Age, ^{14}C Residence Time and "Tape Recorder" Tracers

The previous section provides only indirect evidence that aspects of the stratospheric transport circulation have improved in E2.2. Here we explicitly examine the transport circulation through comparisons of the stratospheric mean age (Γ_{STRAT}) and tropical ascent phase estimates inferred from the idealized tape recorder tracer (χ_{tape}). Given that neither the mean age nor tape-recorder tracers are directly observable, we compare the ModelE results with both in-situ based approximations (derived from SF_6, CO_2 and H_2O) as well as the CCMI simulated fields.

We begin with the mean age (Fig. 2a,b), Γ_{STRAT}, which features characteristic isopleths that ascend in the tropics and slope down over higher latitudes in both hemispheres with values in E2.2-AP (E2.1) corresponding to $\sim$1 (0.5) year(s) in the tropical lower stratosphere and $\sim$5 (3.5) years over polar latitudes in the middle stratosphere. Direct comparisons of Γ_{STRAT} at 20 km (Fig. 3a) as well as over the tropics (Fig. 3b) and northern midlatitudes (Fig. 3c) show very good agreement between E2.2-AP and observational estimates, as well as with the CCMI models. (Note that the differences between the three individual ensemble members is very small). While the ages in E2.2 are slightly (5%) younger than in E2.2-AP, those differences are negligible compared to the differences relative to E2.1, for which the mean age values are $\sim$30% too young, consistent with the values reported for previous versions of the lower vertical resolution version of ModelE (Shindell et al., 2013).

In addition to the mean age, Fig. 3d shows changes in the stratospheric concentration of ^{14}C as a function of month after release in the mid-stratosphere, compared among E2.2-AP, E2.2 and E2.1. Results are presented using natural log coordinates, and also shown are the stratospheric residence time, $\tau_{14\mathrm{C}}$ (the inverse of the associated least mean squares slope). While it is well known that $\tau_{14\mathrm{C}}$ evolves as a function of time after the pulse is released (consistent with changes in the stratospheric distribution of that tracer) for sake of brevity we consider here only residence times after 100 months, for which $\tau_{14\mathrm{C}}$ corresponds to 4.76, 4.63 and 4.21 years for E2.2-AP, E2.2 and E2.1 respectively. Comparisons with the observed values presented in Prather and Remsberg (1993) indicate that, overall, the greater stratospheric residence times in E2.2(-AP) are more consistent with observations for this particular radionuclide, a result that complements the mean age assessment presented earlier. Further comparisons of the spatial distribution of ^{14}C (not shown) also show more coherent cross-tropopause transport in the vicinity of the northern subtropical jet in E2.2(-AP), compared to the relatively more noisy pattern exhibited in E2.1.

A similar calculation was made for pre-industrial times, and the residence times were about five months longer for the E2.2 models, and 1 month longer in E2.1. This suggests that the stratospheric residence time in the more modern time period has decreased, with the effect more noticeable in the E2.2 runs. This might well be associated with an increase in the residual circulation due to global warming in the model, an effect less evident in E2.1 due to its cruder representation of that circulation.

While the mean age and ^{14}C-based residence time represent integrated measures of the effects of both mixing and advection on stratospheric transport timescales, the vertical propagation of the tape recorder tracer provides a more direct measure of the advective transport timescale for ascent to occur within the tropics (Figure 2c,d). In particular, we focus on the evolution of χ_{tape} over 5 years at the beginning of the simulations (1980-1985). This is because during the course of the (multi-year-long) simulations the near-tropopause gradients of χ_{tape} weaken substantially, since that tracer is not subject to any stratospheric or tropospheric loss.

Examination of the tape recorder phase lag, ϕ_{tape} (Fig. 3e), shows good agreement between E2.2(-AP) and observational estimates derived from HALOE water vapor measurements. While ascent is slightly faster in E2.2 compared to E2.2-AP, consistent with the slightly younger mean ages in that model version, this difference is smaller than the differences relative to E2.1, for which the phase lag values are consistently ∼25% too small, compared to the observations. Similar differences in tropical ascent are evident in comparisons of the tape recorder inferred directly from simulated water vapor (Supplementary Figure 3). Note that in Fig. 3e the phase lags from the CCMI models are not shown as they did not integrate the tape-recorder tracer.

Finally, in order to provide a related, but distinct, measure of the strength of in-mixing from the subtropics into the tropical stratosphere, we compare profiles of CH_4, averaged over 10°S-10°N (Fig. 3f). To ensure consistency with the analysis presented in CCMVal (2010) (see their Figure 5.7) we normalize the climatological tropical mean methane profiles from all models (including the CCMI output) to the HALOE values at the bottom level of the region of interest (100 hPa). All model output is then interpolated to the HALOE vertical levels and then compared directly against the observations. Below 30 hPa the methane vertical gradients exhibited in E2.2(-AP) are in line with the observational estimates and the CCMI models, while the vertical gradients in E2.1 are relatively stronger (Fig. 3f). Above 50 hPa the methane vertical gradients are stronger than those exhibited in the observations (but still within the CCMI intermodel spread), which could be indicative of excessive in-mixing into the tropical middle stratosphere, as alluded to earlier. The region over which this occurs, however, is relatively narrow as it is confined to the tropics spanning 10-30 hPa and may explain some of the negative ozone biases in that region exhibited in E2.2(-AP).

3.1.3 Seasonality of Stratospheric CH_4, N_2O, O_3 and H_2O

Comparisons of the seasonal cycle of CH_4, N_2O, and H_2O, averaged over the middle-to-lower stratosphere (30-100 hPa), also show overall good agreement between E2.2(-AP) and the observations, albeit with some biases depending on the latitude (Fig. 4a-b,d, Supplementary Fig. 2a-b,d). Compared to E2.1, E2.2-AP exhibits stronger meridional gradients in the subtropics along with their seasonal migration, consistent with a stronger subtropical transport barrier in that model. This is especially evident in CH_4 (Fig. 4a), H_2O (Fig. 4d, Supplementary Fig. 2d) and in N_2O (Fig. 4b, Supplementary Fig. 2b).

The seasonal cycle of total column ozone in E2.2-AP also compares well overall with the TOMS/OMI observations, with maximum values of ∼415 DU over the NH pole during boreal winter, compared to larger values (∼450 DU) in E2.1 (Fig. 4c, Supplementary Fig. 2c). Over the tropics both models exhibit low ozone biases compared to the observations, albeit smaller ones in E2.2-AP (5%) compared to E2.1 (∼10%). The smaller low tropical and high polar ozone biases in E2.2-AP are most likely related to a weakening of the Brewer-Dobson circulation, compared to E2.1, as reflected in the older mean ages in that model, with associated weaker transport of ozone-rich air from the tropics to high latitudes. Over the SH high latitudes both E2.1 and E2.2 overestimate observed values during austral winter by about 20-30%, although these biases are slightly smaller in E2.2-AP. While the latter is most likely also associated with an overall weaker circulation in E2.2(-AP) it is also likely that the somewhat improved SH ozone burdens poleward of 60°S may reflect the improved lower stratospheric temperatures over the SH pole. In particular, the warm austral winter temperature biases in E2.1, which exceed 20 degrees (see Fig. 3c in R20), are reduced to 2-4 degrees in E2.2(-AP) (see Fig. 2b, also in R20).

Finally, it is important to note that the ozone hole in E2.2-AP does not extend quite long enough during September-October-November. Comparisons with the seasonal cycle of H_2O (Fig. 4d) and CH_4 (Fig. 4a), indicates that this may be driven by an early bias in the seasonality of methane-oxidation driven water vapor production, which may impact heterogeneous ozone depletion on polar stratospheric clouds (PSC).

3.1.4 Interannual Variability Associated with the Quasi-Biennial Oscillation

One of the key dynamical features introduced in E2.2 that distinguishes it from other GISS models is an accurate interactively generated Quasi-Biennial Oscillation as described in R20. The QBO is a dominant mode of transport variability in the stratosphere that impacts a broad range of trace gases including H_2O, hydrogen chloride (HCl), O_3, N_2O, carbon monoxide (CO), hydrogen fluouride (HF), and CH_4 (Schoeberl et al., 2008; Tweedy et al., 2017), with important implications for the detection of lower stratospheric ozone recovery (Chipperfield et al., 2018), among other applications.

The evolution of the tropical (5°S-5°N) zonal winds for one ensemble member of E2.2-AP (Fig. 5a) compares well with observations (R20) and compared to other CMIP6 models (Orbe et al., 2020), albeit with an amplitude that is about 15% less than observed. (Note that the QBO period difference is negligible between the AMIP configurations of E2.2 (28.5 months) and E2.2-AP (27.7 months) considered here). The imprint of the QBO is manifest in E2.2-AP on a broad range of both chemical and idealized tracers, including methane (Fig. 5b), the stratospheric mean age (Fig. 5c) as well as ozone (Fig. 5d).

In particular, below 10 hPa all species exhibit a clear downward propagation of anomalous negative (positive) values for CH_4 (Γ_{STRAT}, O_3) during the westerly shear phase of the QBO, and vice versa. The fact that the anomalies in CH_4 are anti-correlated with those in the mean age and ozone is consistent with the opposite vertical gradients in that tracer (Fig. 1). In particular, during the westerly phase of the QBO the anomalous downwelling associated with warmer anomalies in the tropics draws larger values of Γ_{STRAT} and O_3 and smaller values of CH_4 into the lower stratosphere. Conversely, easterly wind shear requires colder tropical temperatures and the associated upwelling anomalies bring air from the troposphere into the lower stratosphere, thereby reducing the age of air (and ozone).

While the signature of a QBO extends coherently above 10 hPa to 3 hPa for both methane and the mean age, the ozone anomalies display a more complicated relationship, indicative of a transition between photochemically- versus dynamically-driven regimes (Rind et al., 2014). In particular, above 10 hPa the variability in tropical ozone is shorter and controlled more directly by variations in photolysis; furthermore, in addition to directly affecting the transport of ozone, the warmer temperatures associated with the westerly phase of the QBO at these levels also drive less ozone production. By comparison, below 10 hPa the ozone variations become more clearly modulated by transport associated with QBO dynamics and more closely mirror those of the other tracers. Note that, while the amplitude of the QBO in E2.2 compares well with observations at ~60 hPa (Rind et al., 2020), below 70 hPa the amplitude is weaker than observed, consistent with biases exhibited in other models that produce a QBO (Bushell et al., 2020). In addition to potentially limiting the representation of QBO teleconnections to higher latitudes, this bias also likely affects the amplitude of QBO-driven trace gas variability in the model.

3.1.5 Historical Ozone Changes over 1960-2014

While the focus of the previous section was on the validation of key stratospheric species with the available satellite measurements, here we consider how the E2.2-G AMIP historical simulations reproduce the evolution of ozone. Figure 6a shows the evolution of global mean total column ozone in E2.2(-AP) compared with E2.1 and the CCMI models and against both the ground-based measurements from 1960-2014 (Fioletov et al., 2002) and the SBUV v8.6 merged dataset extending back to 1970 (R. D. McPeters et al., 2013). Overall, the total global (90°N-90°S) ozone column decreases over the 1970-1990s are well captured in both models, although the E2.2-AP values are generally smaller than observed; both models also simulate ozone recovery during the years following stratospheric ozone depletion, albeit with lower values. (Note that, while there are some differences in variability during the 1970s, the ground-based and satellite-derived ozone fields agree very well in the global mean). This bias aside, however, the overall performance in both E2.2(-AP) and E2.1 is well within the range of the CCMI models and represents a significant improvement in total column ozone relative to versions of ModelE prior to CMIP5, as discussed in Shindell et al. (2013) and in Kelley et al. (2019).

Comparisons of the Southern Hemisphere (SH) (Fig. 6b) and NH (Fig. 6c) mean total column ozone values also show similar behavior between E2.2(-AP) and E2.1, with somewhat lower values in the former. The overall long-term behavior is similar as well, except that E2.2(-AP) appears to simulate a larger ozone response to the eruption of Mount Pinatubo as well as a faster signature of ozone recovery during the 2000s, which may be linked to the former. In addition, over the Southern Hemisphere there is a distinct discrepancy in the temporal evolution of ozone over the second half of the 1960s between the models. Further inspection of the full historical period (not shown) reveals that these differences in ozone variability are related to differences in the evolution of nitric acid (HNO_3) following volcanic eruptions, with HNO_3 increasing in E2.2 but decreasing in E2.1. While the response in the former is more consistent with expectations, it is not clear why the models diverge and further work is needed to understand how layer-dependent assumptions made within the code that translate (prescribed) aerosol optical depth to aerosol size may be driving these differences. Finally, there is some indication of larger interannual ozone variability in E2.2-AP over the Northern Hemisphere, relative to E2.1. How (if) this is linked to the more realistic polar vortex variability in that model (R20, Ayarzagüena et al. (2020)) will be examined in future studies.

3.2 Tropospheric Transport Circulation

Since the main focus of E2.2 development was towards optimizing the representation of the middle atmosphere, our exposition of the transport characteristics of the troposphere is relatively briefer, compared to the previous section. In addition, there are relatively fewer direct observable constraints on tropospheric transport, owing to the more complex source/sink and emissions distributions of tracers in the troposphere, compared to the stratosphere. Therefore, with the exception of SF_6, which we use to constrain interhemispheric transport, our focus is primarily on the idealized tracers presented in Orbe et al. (2019) and comparisons with the CCMI models.

3.2.1 Transport to the Arctic

The NH midlatitude loss tracers $\chi_{\mathrm{NH},5}$ (Fig. 7a) and $\chi_{\mathrm{NH},50}$ (Fig. 7b) show overall good agreement between the ModelE simulated fields over northern midlatitudes and are within the range of the CCMI models, albeit with ~10% larger values in E2.1, compared to those in E2.2(-AP); by contrast, poleward of 40°N both E2.2(-AP) and E2.1 are biased high, especially for $\chi_{\mathrm{NH},50}$. (Note that the E2.2 versus E2.2-AP differences are negligible). While the high bias in $\chi_{\mathrm{NH},50}$ exhibited in both E2.1 and E2.2 is no larger than ~10% compared to the CCMI multi-model mean, it is nonetheless consistent with the large carbon monoxide Arctic burdens reported for the CMIP5 version of ModelE by Lee et al. (2013), as compared to models participating in the Atmospheric Chemistry and Climate Model Intercomparison Project (ACCMIP). It is also consistent with the high tropospheric ozone column biases over northern high latitudes noted in Shindell et al. (2013) for previous versions of ModelE (see their Figure 3c). This suggests that this transport bias, which may have implications for black carbon loading over the Arctic (and associated radiative forcing), may reflect a systematic, albeit relatively small, bias in the model.

To explore potential dynamical drivers of the high $\chi_{\rm NH,50}$ bias over northern polar latitudes we perform a seasonal decomposition of $\chi_{\rm NH,50}$ and find that the higher tracer burdens in E2.2(-AP) and E2.1 occur primarily during boreal winter (Supplementary Figure 4a,b). Furthermore, comparisons with the climatological mean zonal winds (bottom panels) indicates that the biases in $\chi_{\rm NH,50}$ are associated with biases in the northern midlatitude jet, which is systematically shifted upward in all versions of ModelE, relative to the CCMI models (Supplementary Figure 4c). More precisely, as the midlatitude jet represents a mixing barrier to along-isentropic poleward transport of surface (high $\chi_{\rm NH,50}$) tracers, its upward shifted bias would result in more vigorous along-isentropic transport to the high latitude upper troposphere (see schematic in Supplementary Figure 5). Note that this argument is primarily associated with transport during boreal winter, when along-isentropic mixing provides the dominant mechanism for transport from midlatitudes to the polar region (Klonecki et al., 2003; Orbe et al., 2013). During other seasons such as JJA the mean meridional circulation and across-isentropic transport driven by convection become more important (Hess, 2005; Yang et al., 2019).

3.2.2 Interhemispheric Transport

We also briefly evaluate simulated interhemispheric transport in terms of the mean age since air was last at the NH midlatitude surface, $\Gamma_{\rm NHMID}$ (Fig. 7c). The meridional profile of $\Gamma_{\rm NHMID}$, as simulated in E2.2(-AP) and E2.1, increases sharply from values of $\sim$0.5 years in the northern tropics to $\sim$1.8 years over SH midlatitudes. All ModelE simulations (E2.1, E2.2-AP, E2.2) exhibit similar age profiles and fall well within the spread spanned by the CCMI models. Furthermore, comparisons with observed SF_6 age (a_{SF6}) profiles (equation (2)), reveals that E2.2 (and E2.1) exhibits mean ages that are older than observed, consistent with the CCMI models.

To enable a fairer apples-to-apples comparison between $\Gamma_{\rm NHMID}$ in the models and a_{SF6} derived from the observations we also show a_{SF6} for E2.2-AP, as calculated from simulated SF_6 concentrations (cyan line, Fig. 7c). The close correspondence between a_{SF6} and $\Gamma_{\rm NHMID}$ in the SH for E2.2 demonstrates that the mean- and SF_6- based ages agree well over latitudes far enough away from the (northern midlatitude) source region. (Note that we only show $\Gamma_{\rm NHMID}$ for the CCMI models as significantly more models integrated that tracer, compared to only two models which integrated SF_6). The old age bias in the ModelE simulations, manifest in both a_{SF6} and $\Gamma_{\rm NHMID}$, is consistent with a similar bias in the CCMI models (Orbe et al., 2019) as well as the TransCom chemistry transport (offline) models (Yang et al., 2019). While the latter study posits that the age biases are likely driven by biases in transport processes between the northern tropics and extratropics, a more in-depth examination of the biases in ModelE is beyond the scope of this study and is reserved for future work.

3.2.3 Upper Troposphere/Lower Stratosphere Transport

As in Abalos et al. (2017) we use the idealized loss tracer e90 to evaluate transport in the upper troposphere and lower stratosphere (UTLS). While e90 does not provide a direct estimate of the stratosphere-to-troposphere (STT) (or troposphere-to-stratosphere) air mass *flux*, versus other measures (Appenzeller et al., 1996; Gettelman & Sobel, 2000; Orbe et al., 2012), it correlates well with stratospherically sourced idealized tracers (i.e. $\chi_{\rm ST8025}$ analyzed in Orbe et al. (2019)) and, most importantly, was incorporated in the CCMI idealized tracer package, thus providing a means for comparing UTLS transport properties of E2.2 with those of other models.

As shown in Fig. 7d there is good agreement in the meridional 100-500 hPa averaged distributions of e90 between E2.1, E2.2-AP and E2.2 and with the CCMI models over the subtropics and tropics. Poleward of 40°S/N, especially in the NH, all versions of ModelE tend to exhibit larger values, compared to the range spanned by CCMI. The e90 biases are consistent in amplitude (and order among the ModelE simulations) with the biases in $\chi_{NH,50}$, which is somewhat to be expected given that both tracers have surface sources. However, given its longer lifetime and more global surface source distribution, we also expect that the e90 biases not only reflect potentially excessive transport from the NH midlatitude surface into the northern upper troposphere, but also other transport processes occurring in the lower stratosphere.

In particular, comparisons of the ensemble mean annually and hemispherically averaged climatological cross-tropopause flux over the NH for E2.1 (-2.5$\times 10^{-4}$ kg/m^2/s) compared to E2.2-AP (-4.3$\times 10^{-4}$ kg/m^2/s) (and -4.1$\times 10^{-4}$ kg/m^2/s for E2.2) reveal weaker downward mass transport in E2.1, consistent with less downward exchange of low-e90 stratospheric air masses into the upper troposphere. This suggests that differences in transport from the stratosphere may also be playing a role in the e90 biases over high latitudes. It is important to note that both the cross-tropopause flux and e90 tracer signatures are not merely reflections of circulation differences over high latitudes but also differences in tropopause height among the models, which is biased high over the tropics and northern extra-tropics in E2.1, compared to E2.2(-AP) and the CCMI models (Supplementary Figure 6). Namely, higher mid-to-upper tropospheric burdens of e90 are consistent with a higher tropopause in E2.1; by comparison, both the tropopause and e90 values in E2.2(-AP) are in better agreement with the CCMI models (Fig. 7d), particularly equatorward of 40°S/N.

3.2.4 Vertical Transport

One contribution to the interhemispheric transport biases in the model could, among other factors, be related to differences in convective transport in the northern tropics and subtropics. In particular, Orbe et al. (2019) showed that the Γ_{NHMID} differences among the CCMI models were well correlated with the strength of northern subtropical convection. While their focus was on the strength of (parameterized) convection over the subtropical *oceans*, it is nonetheless still useful to also examine the simulated distributions of the radon tracer, despite the fact that it can only be credibly validated over land, owing to the limited available observations.

Given its short life-time (half-life of 3.8 days), Rn-222 is most responsive to rapid transport, such as that associated with convection. However, in evaluating the factors shaping the Rn-222 distribution one has to consider the consequences of convection on the circulation as well the direct impact of convective transport itself. By rapidly redistributing heat and momentum, convection induces meridional and zonal circulations which can then also advect radon. At low elevations, turbulence/dry convection act to redistribute radon away from its surface source.

Five-year average model profiles of Rn-222 are shown in Figure 7e for a location coincident with a set of two observations (12°N, 76°E). In the low and mid troposphere E2.1 exhibits larger values than the two E2.2 models which produce equivalent results. Analysis shows that this is due to greater gain by turbulence/dry convection and meridional transport in E2.1, which more than balance the greater loss by convective transport. At the same time, this greater convective transport promotes increased values above 4km compared with the other models. It is also worth noting that the (high) bias in E2.1 at this latitude occurs in exactly the same altitude range where E2.1 exhibits excessive water vapor (R20, Figure 33) and stronger moist convective mass fluxes, compared to E2.2(-AP) (R20, Figure 35). The two different observed profiles (Rani et al., 2014), one from early in the morning and one early in the afternoon are also shown in Fig. 7e (all results are normalized to be equal to 1 at the surface). Given the disparity in data sample sizes, however, no meaningful comparison is possible, although taken at face value the E2.2 models produce profiles whose time-average is similar to that shown in the observed profile. Clearly, more tropical and subtropical Rn-222 observations are needed to more credibly evaluate model convection and circulation in these regions.

A comparison of these models with Rn-222 observations was also made over northern midlatitudes (Murray et al. (2014), shown in Supplementary Figure 7); the observations are averages of three or more data retrievals for each month. The results do not indicate a consistent model high bias at either low or high elevations, with model values exceeding observations in the upper troposphere in some months, while the reverse is true in others. The Rn-222 distribution above 8 km at these latitudes results from gain by convection and resolved vertical advection, and loss by meridional transport. From month-to-month the balance between these processes changes; thus, for example, the models have large values in the upper troposphere in July associated with large gains from convection and resolved vertical advection, and in September due to less loss from meridional transport. In August smaller values occur, from greater meridional transport loss and smaller convective and resolved vertical transport gains. Ultimately these variations result from the monthly/model differences in convection, but the complexity of this diagnosis is partly due to having convection in its own subroutine outside of the model's numerical solution of the conservation equations, resulting in air mass and tracer transport fragmented between the two different subroutines.

Finally, Orbe et al. (2019) showed that convection over the subtropical oceans is associated with both less efficient transport to high northern latitudes and faster interhemispheric transport to high southern latitudes. The fact that the transports to high latitudes are very similar between E2.1 and E2.2(-AP) therefore indicates that the (small) differences in convection over land do not influence interhemispheric transport in E2.1/E2.2.

4 Dynamical and Transport Responses to Idealized Increases in CO_2

Having validated the large-scale transport characteristics of E2.2(-AP) in the previous sections next we document changes in the stratospheric and tropospheric transport circulations in response to both abrupt and transient increases in CO_2 through use of the CMIP6 DECK experiments (Table 1). Following Grise and Polvani (2016) we term the changes in the dynamical circulation the "dynamical sensitivity" of the model and, by extension, those in the transport circulation the "transport sensitivity."

The dynamical responses in the E2.2(-AP) CO_2 experiments were discussed only briefly in R20, specifically with regards to changes in global mean surface temperature and the response of the North Atlantic meridional overturning circulation. The only study examining the stratospheric response so far (Ayarzagüena et al., 2020) only did so with respect to stratospheric polar vortex variability (i.e. sudden warming events). Therefore, though not exhaustive in our analysis, here we take the opportunity to present basic measures of the stratospheric and tropospheric dynamical responses that were not examined in the previous studies but are important for interpreting the transport response, which is the focus of this study. All responses (δ), which herein are defined as the difference between the last 50 years of the abrupt and transient CO_2 simulations, relative to the PI control, are discussed in the text and summarized accordingly in Table 3.

4.1 Stratospheric Dynamical and Transport Circulations

4.1.1 Residual Mean Circulation (Ψ^*)

The residual mean circulation (Ψ^*), here used to describe the advective component of the Brewer-Dobson circulation, accelerates throughout the stratosphere as CO_2 is increased (Fig. 8, top), a change that is among the more robust dynamical responses in models (Rind et al., 1998, 2002; Butchart & Scaife, 2001; Sigmond et al., 2004; Garcia & Randel, 2008; Li et al., 2008; Oberländer et al., 2013; Butchart et al., 2010; Hardiman et al., 2014). While in more realistic climate change scenarios changes in ozone depleting substances may significantly weaken this projected acceleration over the 21st century (Polvani et al., 2018; Abalos et al., 2019) it is nonetheless important that the underlying CO_2-induced response of the BDC be rigorously understood.

The changes in the residual mean circulation manifest in the lower stratosphere as an increase in annual mean upwelling (ω^*) throughout the tropics and subtropics ($\sim$ 20°S-20°N) (Fig. 8, bottom). In particular, lower stratospheric upwelling increases by 17±7.4%, 16±8.8% and 39±9.7% for 1%CO_2, 2xCO_2, and 4xCO_2, respectively (Table 3, row 2). At higher altitudes (10 hPa) the changes $\delta\omega^*$ are much smaller, at 5.4±13 % (1%CO_2), 4.2±12 % (2xCO_2), and 8.4±13% (4xCO_2). Overall, the good correspondence between the changes in Ψ^* (ω^*) between the 1%CO_2 and 2xCO_2 experiments indicates that the (equilibrated) acceleration of the circulation is not sensitive to the functional form of the prescribed CO_2 forcing (Fig. 8 left two panels, top and bottom). Note that 2xCO_2 levels are reached around year 70 of the 1%CO_2 integration. Furthermore, the fact that the 4xCO_2 response is nearly double that of the response in the 2xCO_2 experiment throughout the stratosphere indicates that the circulation scales approximately linearly with CO_2, at least within the radiative forcing spanned by the range of CO_2 perturbations considered in this study.

Comparison of the fully coupled 2xCO_2 and 4xCO_2 experiments with those in the FIXSST experiments indicates the extent to which the residual circulation changes reflect rapid adjustments versus changes induced by SST warming. In the lower stratosphere (70 hPa), the response in upwelling is dominated by feedbacks associated with changes in SSTs, as indicated by the changes in $\delta\omega^*$ for the fixed-SST simulations, which are equal to only 2.4% and 5.2% for 2xCO_2 and 4xCO_2, respectively, compared to 16% and 39% for the fully coupled simulations (Table 3, row 2). By comparison, in the middle stratosphere (10 hPa) the values $\delta\omega^*$ for the fixed-SST runs (3.9% (2xCO_2) and 7.8% (4xCO_2)) are much more comparable to the upwelling response in the fully coupled system (5.4% (2xCO_2) and 8.4% (4xCO_2)). Therefore, in the middle stratosphere rapid adjustments contribute significantly more to the circulation response compared to in the lower stratosphere, where SST changes play a key role in modulating the full coupled response.

4.1.2 Stratospheric Transport Circulation ($\delta\Gamma_{\mathrm{STRAT}}$ and δe90)

The response of the transport circulation is consistent with increases in the residual mean circulation (Fig. 9). Specifically, the mean age of air (Fig. 9, top) decreases throughout the stratosphere; at the same time, e90 (Fig. 9, bottom) increases throughout the tropical lower and middle stratosphere and over the midlatitude lower stratosphere, with large positive anomalies straddling the tropopause. The responses in both tracers are qualitatively consistent with Butchart and Scaife (2001), Butchart et al. (2010), and Oman et al. (2009), for the case of the mean age, and Abalos et al. (2017) (see their Figure 8), for the case of e90, albeit for the more comprehensive forcings considered in those studies.

Over the tropics the transport changes are consistent in amplitude with the increases in upwelling (ω^*), with values of Γ_{STRAT} in the tropical lower stratosphere decreasing by 21±8.3%, 22±8.2%, and 40±7.5 % for 1%CO_2, 2xCO_2, and 4xCO_2, respectively (Table 3, row3). At the same time e90 increases by 26±9.7% (1%CO_2), 26±11% (2xCO_2) and 54±12% (4xCO_2) (Table 3, row 4). In addition to being consistent in magnitude with the changes in upwelling (Table 3, row 2), the transport responses in the tropics are also generally linear with CO_2 (Table 3, column 4 vs. 5) and insensitive to the transient nature of the forcing (Table 3, column 2 vs. 3).

Over higher latitudes in the lower stratosphere, the behavior of the mean age and e90 tracers are also linear across the 2x- and 4xCO_2 experiments (Table 3, rows 5-6), although a bit weaker in amplitude compared to the tropics, with Γ_{STRAT} decreasing by 24±8.1% poleward of 60°N (compared to 40% in the tropics) for the 4xCO_2 simulation (Table 3, rows 5) and e90 poleward of 60°N increasing by 32±8.9 %, compared to 54±12% in the tropics (also for 4xCO_2). This weaker response at high latitudes reflects the fact that the extra-tropical transport properties are not solely controlled by changes in tropical advection, but also by mixing between the tropics and extra-tropics, with enhanced mixing resulting in older mean ages over high latitudes (Neu & Plumb, 1999). In addition, as shown in Abalos et al. (2019) the high latitude changes in lower stratospheric e90 are to a large extent tethered to changes in tropopause height, which rises in response to increased CO_2, as discussed further in the next section. Therefore, both changes in extra-tropical mixing and tropopause height likely modulate the amplitude of the high latitude e90 response to increased CO_2.

Interestingly, while the mean state of the stratospheric transport circulation scales more or less linearly with CO_2, the transport variability changes nonlinearly, especially in the tropics (Fig. 10). In particular, the amplitude of interannual variability in Γ_{STRAT} in the 4xCO_2 simulation (shown after the tracer response has more-or-less equilibrated) underestimates that predicted by linearity by more than 50% in the tropics (Fig. 10b) and 25% over high latitudes (Fig. 10a,c). This is consistent with a weakening in the amplitude of the Quasi-Biennial Oscillation as CO_2 increases (Supplementary Figure 8). Note that, unlike changes in the period of the QBO, the weakening in QBO amplitude is a robust response among models as CO_2 increases, as documented in Richter et al. (2019) for models participating in the SPARC Quasi-Biennial Oscillation initiative (QBOi), although the implications for transport were not explored in that study. While our presentation here has been brief in keeping with the broad scope of this study, future work will focus on further disentangling the role of the QBO in E2.2 on simulated transport variability in the stratosphere and its response to climate change.

Finally, we exploit the abruptness of the CO_2 forcings in the 2xCO_2 experiment to glean insight into the relationship between the changes in upwelling over the tropics and the responses of Γ_{STRAT} (and e90) (Fig. 11). In particular, the evolution of global mean Γ_{STRAT} (e90) at 70 hPa is shown to negatively (positively) covary closely with lower stratospheric tropical upwelling (correlations > 0.8, Fig. 11 a,b). Furthermore, the evolution of the upwelling and transport responses are observed to hold not only on interannual timescales but also in the initial SST-mediated response to CO_2 that occurs within the first 10-15 years after the forcing is applied. Given the important role of SST changes in the upwelling response, illustrated earlier through use of the fixed-SST experiments, we also find that the tropical upwelling responsible for the transport changes are strongly correlated with the changes in tropical SST warming (Fig. 11c).

4.2 Tropospheric Dynamical and Transport Circulations

4.2.1 Tropospheric Dynamical Circulation ($\delta\Psi^*$, $\delta \mathrm{p_{trop}}$ and δMCFLX)

First we consider the changes in the residual mean streamfunction in the troposphere (Fig. 12, top). (Note we do not use more conventional (Eulerian) measures for the mean meridional circulation (MMC) because, as in the stratosphere, our discussion is oriented around changes in the transport circulation, for which the residual mean circulation is most relevant). Overall, the changes in Ψ^* reflect a narrowing and acceleration in the mid-to-upper tropical troposphere in agreement with Li et al. (2010), accompanied by a deceleration and weakening throughout the midlatitude troposphere in both hemispheres. As with the stratospheric transport and dynamical changes, the changes in the abrupt 2xCO_2 integration are very similar and almost identical to those produced in the transient forcing run.

Unlike in the stratosphere, comparisons of (double) the 2xCO_2 changes in $\delta\Psi^*$ with those from 4xCO_2 reveal much weaker linearity in both hemispheres. This is generally consistent with Marvel et al. (2015) who show that precipitation responds nonlinearly in ModelE to different anthropogenic forcings, albeit for a broad range of trace gas and aerosol forcings used in that study (as opposed to the incremental CO_2 increases considered here). In particular, we find that in the NH (SH) $\delta\Psi^*$ decreases in the 4xCO_2 simulation by 4.9±4.6% (9.2±4.6%), compared to the 14% (5.1%) predicted by linearity (Table 3 row 7). Interestingly, the nonlinearity in $\delta\Psi^*$ is more pronounced over the extra-tropics, compared to the tropics, where one may expect that abrupt changes in convective instability may drive nonlinear behavior in the overturning tropical circulation. Rather, over the extra-tropics the nonlinear behavior in $\delta\Psi^*$ may be related to non-monotonic changes in baroclinic eddies with increasing surface air temperature, as has been explored in O'Gorman and Schneider (2008), albeit using an idealized model.

In particular, further inspection of the baroclinic eddy kinetic energy generation reveals significant nonlinear (and non-monotonic) behavior, especially over the NH, with hemispherically averaged values at 500 hPa decreasing from 53.1×10^{-4} W/m^2/hPa in the PI control to 52.0×10^{-4} W/m^2/hPa in the 2×CO_2 experiment and then increasing to 54.5×10^{-4} W/m^2/hPa for 4×CO_2. Though beyond the scope of this study, future work will focus on exploring nonlinearities in the midlatitude eddy-driven circulation more rigorously through use of a broader suite of CO_2 forcing experiments spanning 0.5-to-8×CO_2.

As the mean meridional circulation weakens and expands the tropopause rises in response to increased CO_2 (Fig. 12, black lines) (Lorenz & DeWeaver, 2007; Lu et al., 2008). Over the tropics the tropopause rises by 6.3±1.8 %, and 6.3±2.0% and 13±2.1% for the 1%CO_2, 2xCO_2 and 4xCO_2 experiments, respectively; the extra-tropical response (δp_{trop}) is similar in magnitude, if not slightly weaker (Table 3, rows 7-8). The tropopause height changes over both the tropics and extra-tropics scale approximately linearly with CO_2 as do the (parameterized) convective mass flux changes δMCFLX (Fig. 12, bottom), which decrease throughout most of the troposphere (excluding the upper troposphere and the Arctic), as predicted by theoretical constraints on the mass exchange between the boundary layer and free troposphere (Held & Soden, 2006). (Note that the latter changes are primarily linked not to a reduction in the zonal-mean overturning (i.e. the Hadley circulation) but, rather, a reduction in the zonally asymmetric Walker circulation (Vecchi & Soden, 2007)). In the next section we discuss what the different responses between the MMC, tropopause height and mass flux measures imply for the tropospheric transport circulation responses to increased CO_2.

4.2.2 Tropospheric Transport Circulation ($\delta\chi_{NH,5/50}$ and $\delta\Gamma_{NHMID}$)

In response to a doubling of CO_2 the changes in the loss tracers $\chi_{NH,5}$ (Fig. 13 top) and $\chi_{NH,50}$ (Fig. 13 middle) consist primarily of weakly negative anomalies (5-10%) throughout the troposphere and a band of positive anomalies at the tropopause. The pattern and magnitude of the response of $\chi_{NH,50}$ strongly resembles that of the response of an idealized air-mass origin tracer presented in Orbe et al. (2015) (see their Figure 3) and the idealized tracer $\chi_{CO,50}$ analyzed among the ACCMIP models in Doherty et al. (2017), albeit for the different scenarios (Ref A1b and RCP 6.0, respectively) considered in those studies. Furthermore, while the increased burdens of the surface loss tracers at the tropopause could be interpreted merely as reflections of increased tropopause height, Doherty et al. (2017) (see their Figure 11) showed that the tracer increases persist even after replotting in "tropopause-relative" coordinates. This suggests that the CO_2 induced changes, $\delta\chi_{NH,5}$ and $\delta\chi_{NH,50}$, are not simply reflections of tropopause height changes but, rather, also reflect changes in along-isentropic transport in the troposphere to high latitudes, as discussed in Orbe et al. (2015).

Comparisons of $\delta\chi_{\mathrm{NH},5/50}$ for the 2xCO_2 and 4xCO_2 experiments indicates that the transport response in the troposphere is nonlinear with increased CO_2. In particular, for $\chi_{\mathrm{NH},5}$ the response to 4xCO_2 is significantly weaker (-7.1±2.2%) than twice the response to a doubling of CO_2 (-11%) (Table 3, row 11). This behavior is much more consistent with the nonlinear behavior in the MMC ($\delta\Psi^*$), compared to the linear changes in extra-tropical convective mass fluxes and tropopause height discussed earlier. Therefore, while vertical mass flux changes have been invoked in previous studies to qualitatively interpret the transport response to climate change (Fang et al., 2011; Doherty et al., 2017) our results suggest that the driving mechanism more likely involves a weakening of the (resolved) mean meridional circulation.

Finally, the response of Γ_{NHMID} to a doubling of CO_2 consists of small (∼5%) primarily positive anomalies throughout the Southern Hemisphere (Fig. 13, bottom). (Note that the decreases in Γ_{NHMID} above the tropopause mirror those for Γ_{STRAT}, discussed in the previous section). The fact that Γ_{NHMID} increases is consistent with an overall weakening of the MMC and with Holzer and Boer (2001), who showed that interhemispheric exchange times, mixing times, and mean transit times all increase by 10% in response to a doubling of CO_2. The fact that the response in E2.2-AP is weaker is likely not related to our use of Γ_{NHMID} as a measure of IHT, given that Holzer and Boer (2001) also used similar integrated measures derived from the age spectrum (Hall & Plumb, 1994; Holzer & Hall, 2000). Rather, the differences most likely reflect differences in the underlying models, particularly with respect to resolution as well as their sea surface temperature response to increasing CO_2; both of these affect the simulated transport sensitivity (Rind et al., 2002). Though beyond the scope of this current study, future work will focus on further understanding how the transport sensitivity in E2.2 varies with resolution, choice of convective parameterization, coupling to the ocean and other factors.

5 Sensitivity Analysis

Here we comment on the sensitivity experiments introduced in Section 2 that guided the development of the "Altered Physics (AP)" version of E2.2 (i.e. E2.2-AP). In particular, various aspects of the convective parameterization were altered in order to remove artificial dependence on layer thickness, including changes made to detrainment, conditional instability, repartitioning of precipitation into lofted and detrained fractions, evaporating precipitation, downdrafts and updrafts. Upon incorporation of some of theses changes Rind et al. (2020) showed that E2.2-AP differs from E2.2 in several respects (e.g. convective mass flux, specific humidity, precipitation, standing wave number 1 energy in the stratosphere). Those differences notwithstanding, however, here we have shown that the large-scale transport properties of the two model versions are nonetheless very similar, at least relative to the larger existing differences between E2.2 and E2.1. Next we show that this is because, unlike in previous "Middle Atmosphere" versions of ModelE, E2.2-AP was developed explicitly with transport considerations in mind.

More precisely, E2.2-AP was developed so as to produce not only a credible QBO period as in past versions (Rind et al., 2014) but also a realistic stratospheric mean age of air. As discussed in R20 the QBO period is relatively straightforward to tune by altering the assigned convective phase velocities and convection/momentum flux scaling within the non-orographic gravity wave drag parameterization. By comparison, the stratospheric mean age, as illustrated in the previous sections (see Figure 11), is strongly tethered to the (resolved) upwelling in the tropical lower stratosphere, which in turn is highly dependent on the model's climatological SSTs. Therefore, upon introducing some of the convective parameterization changes in the (atmosphere-ocean) coupled version of E2.2

we found that the associated increases in SSTs resulted in increased lower stratospheric upwelling, which in turn directly affected the stratospheric mean age in the model.

In particular, Fig. 14a shows how applying the proposed convection changes in both coupled (atmosphere-ocean) and AMIP versions of E2.2 results in a vertical redistribution of climatological mean convection to more “top-heavy” profiles (R20). While this is consistent with the original intention of the proposed changes (i.e. to produce a warmer upper tropical troposphere, considering that E2.2 is biased cold in the troposphere), the SSTs in the coupled (atmosphere-ocean) system also adjust accordingly, increasing by ∼2 degrees, with some changes producing larger SST responses (red, blue, cyan) compared to others (green). (Note that, by design, these changes to the convective parameterization produce similar mass flux responses in the AMIP configurations, but with no associated changes in SSTs).

As the SSTs increase in response to the convective parameterization changes, lower stratospheric upwelling also increases in the coupled simulations (Fig. 14b), a feedback that is absent in the AMIP experiments. These increases in lower stratospheric upwelling are driven by an equatorward shift in the subtropical upper tropospheric jet, which increases the equatorward Eliassen-Palm flux convergence ocurring in the lower stratosphere in both hemispheres from 45°S to 45°N (not shown). A similar jet-mediated pathway linking warmer SSTs to increased upwelling has been reported in previous studies including Rind et al. (2002) and Li et al. (2010), among others.

Associated with these changes in the large-scale flow the QBO periods in the coupled simulations also increase (large filled circles, Fig. 14c). While retuning the non-orographic GWD can correct for the QBO period changes (small filled circles) the large-scale flow changes nonetheless persist and are associated with significantly younger stratospheric mean ages, compared to in the AMIP experiments (Fig. 14d). Therefore, in determining which convection changes were to be incorporated in E2.2-AP we decided only on those that produced *both* QBO periods and mean age values consistent with observations (green filled circles, Fig. 14).

The above illustrates two important aspects of the development of E2.2(-AP). The first relates to optimizing the model not only in terms of its middle atmospheric dynamics but also its transport circulation (mean age). The second relates to the critical role played by testing various parameterization settings in both AMIP and coupled atmosphere-ocean configurations. In particular, the latter captures feedbacks between convection, sea surface temperatures and stratospheric upwelling that cannot be gleaned in a prescribed SST framework but are nonetheless key for evaluating the true coupled nature of the model.

6 Conclusions

The main goal of this study has been to evaluate the large-scale transport characteristics of the new climate model GISS-E2.2 that has been specially optimized for the middle atmosphere and whose output has been contributed to the CMIP6 archive. As such it complements the overview presented in Rind et al. (2020), which discussed in detail the model’s underlying structure, parameterization choices (including departures from those chosen in the lower vertical resolution version of ModelE (E2.1)), and a broad range of key dynamical and radiative properties of the model’s climatology. As in that study, in addition to validating the performance of the model over the historical period, we also present its climate response, with the goal of quantifying the “transport sensitivity” of E2.2 through use of the abrupt 4xCO_2 experiment submitted to CMIP6 as well as additional CO_2 varying (i.e. 2xCO_2 and 1%xCO_2) experiments. Finally, we present results from several sensitivity tests in order to illustrate the large-scale dynamical and transport assessments that were used to inform the parameterized convective and non-orographic gravity wave drag settings that were employed in E2.2-AP.

Our analysis of a broad range of transport measures derived from both real chemical as well as idealized tracers shows various improvements in the stratospheric transport circulation in E2.2 compared to previous versions of the model. Most notably, the stratospheric mean age values in E2.2 represent a dramatic improvement over previous model versions in which the mean ages were too young (Shindell et al., 2013) and likely reflect the weaker tropical ascent in the model as well as a more realistic tropical pipe. In addition to these improvements in the mean state of the stratospheric transport circulation, the transport variability associated with the QBO is also captured in the distributions of various tracers (i.e. CH_4, O_3) owing to the incorporation of non-orographic GWD drag sources from convection and shear that are absent in other versions of ModelE.

For sake of completeness we have also presented a (briefer) evaluation of the tropospheric transport climate in E2.2, which shows similar overall characteristics to those of E2.1. While both versions of the model exhibit somewhat too vigorous poleward transport in the Northern Hemisphere, which may be related to a systematic upward bias in the midlatitude jet, the overall properties of the tropospheric transport climate are within the range spanned by other chemistry climate models, particularly those recently participating in CCMI.

Our validation of E2.2 against both observations and the CCMI models indicate that it is a model well equipped for use in understanding not only recent past but also future changes in the transport circulation. To this end, we have also presented an evaluation of the "transport sensitivity" of E2.2 that goes beyond the standard DECK set of CMIP6 integrations by including a 2xCO_2 experiment as well as fixed SST versions of the abrupt 2xCO_2 and 4xCO_2 experiments. From the former we have assessed the linearity of the transport circulation response to CO_2 and from the latter we have quantified the relative importance of rapid adjustments versus feedbacks to both dynamical and transport responses. Our main findings are as follows:

- In response to doubled (quadrupled) CO_2, E2.2 simulates a ∼20% (∼40%) reduction in the stratospheric mean age and increases in e90 in the tropical lower stratosphere, consistent in magnitude and sign with enhanced upwelling in the tropical lower stratosphere.
- Over the entire stratosphere the transport responses in both Γ_{STRAT} and e90 scale approximately linearly with CO_2 and with SST warming in the tropics.
- Increases in lower stratospheric upwelling are driven primarily by SST warming, with rapid adjustments playing a minor role ($<$ 30%). By comparison, rapid adjustments play a much more important role at higher altitudes (10 hPa).
- In the troposphere E2.2 simulates increased burdens of NH midlatitude surface tracers over the Arctic high latitude tropopause, accompanied by decreased burdens over midlatitudes as CO_2 increases. This response is nonlinear with CO_2, consistent with nonlinear changes in the residual mean meridional circulation. By comparison, changes in tropopause height and vertical mass exchange by (parameterized) convection are much more linear.

While the exact magnitude of the transport responses in E2.2 are likely to depend on the specifics of the model (e.g. resolution, convective parameterization) the overall responses in both the stratospheric and tropospheric tracers are consistent in magnitude and pattern with the results from previous studies. In particular, the changes in Γ_{STRAT} and e90 respectively point to an acceleration of the Brewer-Dobson circulation and enhanced stratosphere-troposphere exchange over both the tropics and extra-tropics. In the troposphere the changes in the idealized loss and mean age tracers are indicative of enhanced Arctic burdens of NH midlatitude surface tracers and (weak) reductions in interhemispheric transport.

A novel contribution from this study is that we have explicitly evaluated the linearity of the transport circulation response in both the stratosphere and the troposphere to increased CO_2. This is motivated partly by the results from a recent study by Abalos et al. (2019) who showed strong correlations between the projected changes in stratosphere-troposphere transport and the amplitude of upper tropospheric warming in the CCMI models (see their Figure 2), indicating the potential for using upper troposphere/lower stratosphere transport measures to constrain climate sensitivity. Here, through use of different CO_2 forcing experiments, we have shown that this relationship also exists in E2.2, in which the stratospheric transport circulation response is strongly correlated with the amplitude of surface warming.

Given the broad scope of this study we have only evaluated the transport circulation and sensitivity of E2.2 in the absence of composition feedbacks through use of the non-interactive version of the model. Given that stratospheric ozone feedbacks can play an important role in modulating "dynamical sensitivity" (Chiodo & Polvani, 2017), however, it remains to be seen how the "transport sensitivity" of E2.2 itself depends on how ozone and other constituents evolve as the planet warms. To this end, interactive versions of E2.2 have also been produced for CMIP6 and will be presented in future studies.

7 Data Availability

All E2.1 and E2.2-AP CMIP6 DECK simulations discussed in this study are available through the Earth System Grid Federation (ESGF). In addition, all of the E2.2 sensitivity and fixed-SST experiments can be found at `https://gmao.gsfc.nasa.gov/gmaoftp/corbe/CMIP6/E2-2-G/`.

Acknowledgments

C.O. thanks Krzysztof Wargan for processing and performing quality screening on the MLS Aura N_2O fields that were used in this analysis, Darryn Waugh for providing the in-situ SF_6 age profiles and HALOE tape-recorder phase lag values, and Eric Fleming for help with accessing the ground-based total column ozone measurements. Climate modeling at GISS is supported by the NASA Modeling, Analysis and Prediction program, and resources supporting this work were provided by the NASA High-End Computing (HEC) Program through the NASA Center for Climate Simulation (NCCS) at Goddard Space Flight Center.

[ht]

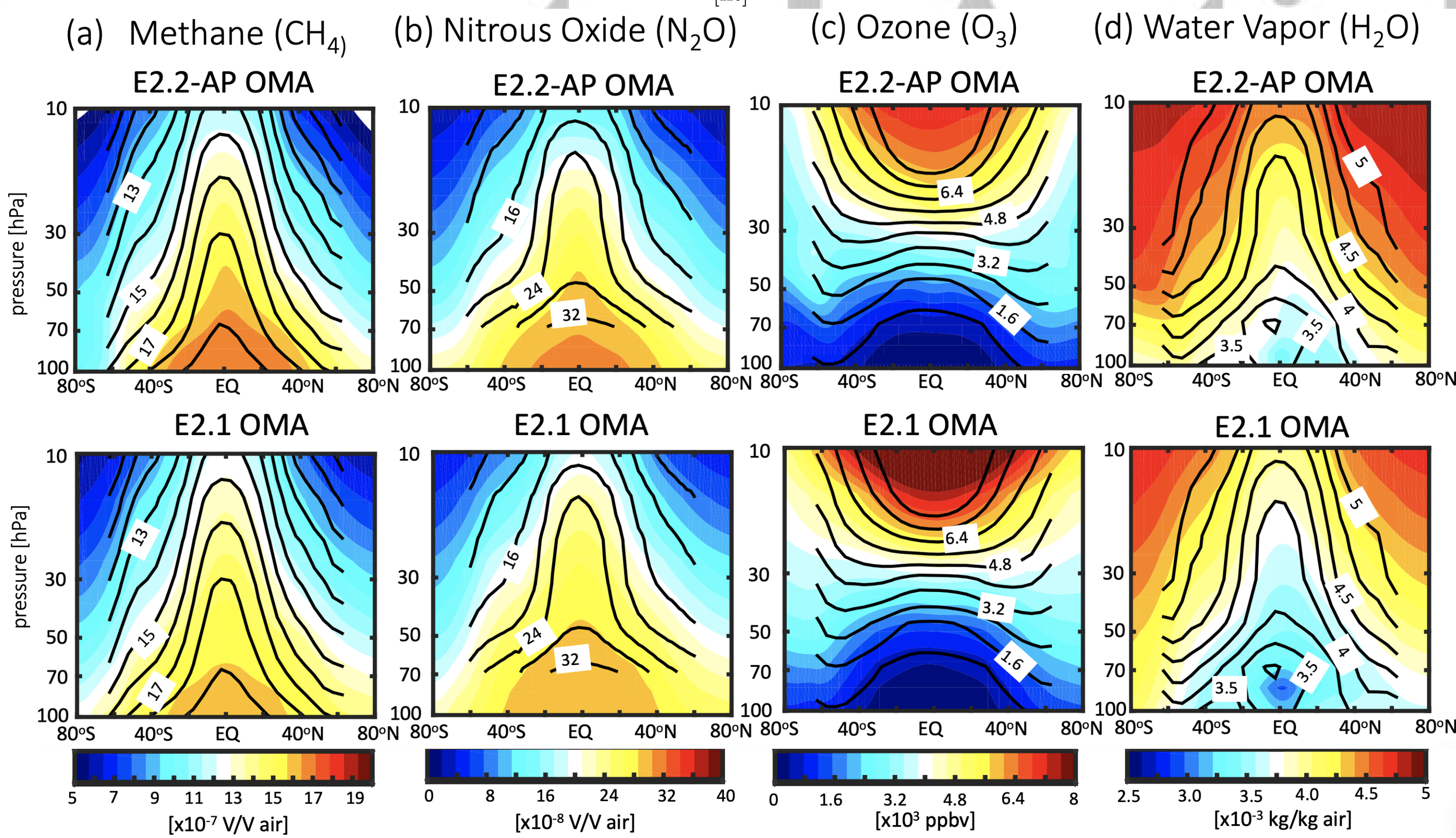

Figure 1. Climatological annually and zonally averaged methane (a) (CH_4), (b) nitrous oxide (N_2O), (c) ozone (O_3) and (d) water vapor (H_2O). Color contours shows the three-member mean of the E2.2-AP OMA Historical ensemble (top) and the two-member mean of the E2.1 OMA Historical ensemble (bottom). Black contours show observed values from HALOE (a,c,d) and MLS (b). Climatologies have been performed over years 1991-2002 when comparing to HALOE (a,c,d) and over 2005-2014 when comparing to MLS (b).

[ht]

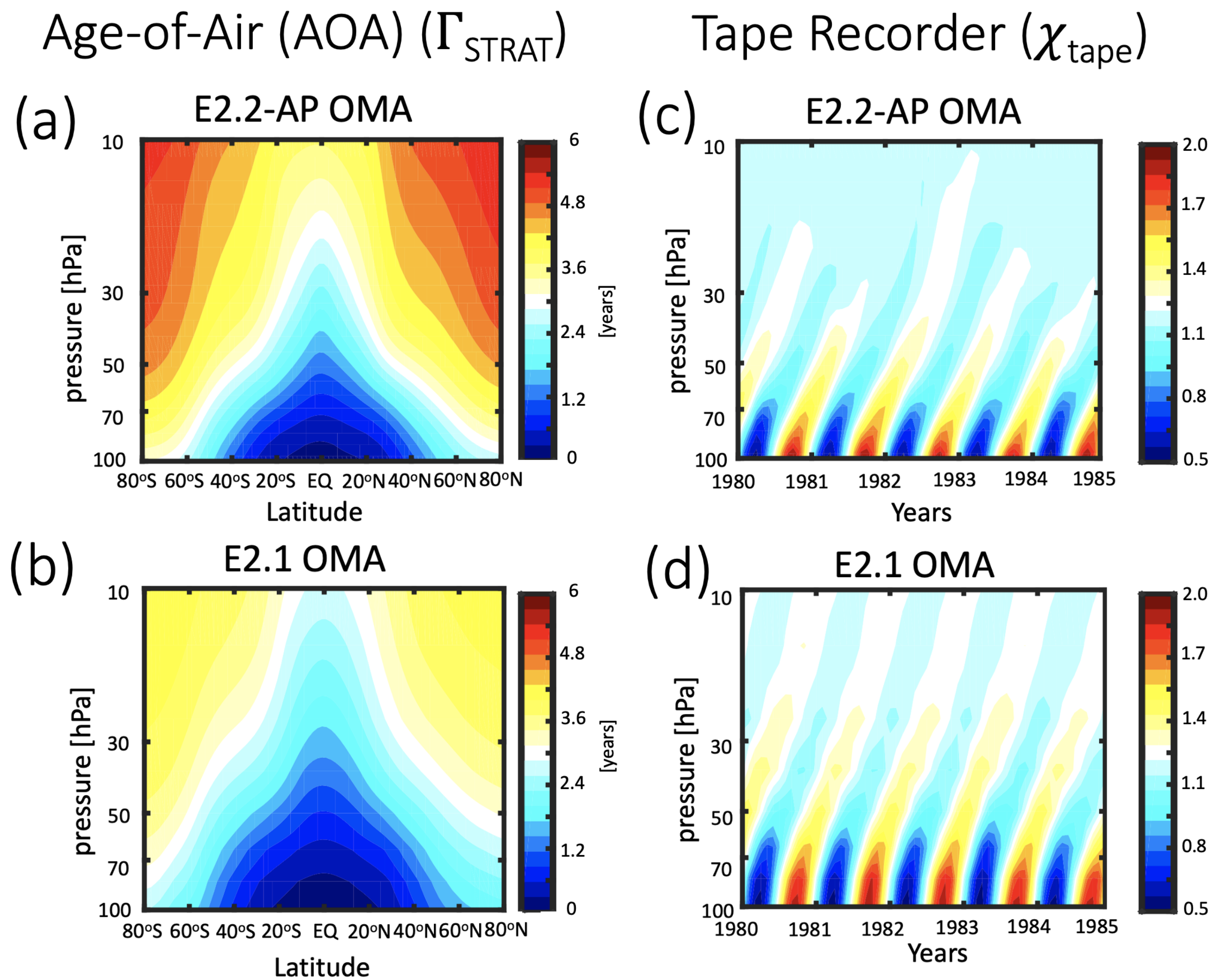

Figure 2. Left: Zonal and annually averaged climatological stratospheric mean age of air (Γ_{STRAT}), averaged over the last five years (2010-2014) of the E2.2-AP (a) and E2.1 (b) OMA Historical ensembles, respectively. Three (two) members are included in the E2.2-AP (E2.1) ensemble means. Right: Evolution of the 10°S-10°N averaged zonal mean tape-recorder tracer (χ_{tape}) for E2.2-AP (c) and E2.1 (d). Note the faster ascent characteristics in E2.1, compared to E2.2-AP. Only one member of E2.1 is shown in (d) in order to avoid averaging over (offset) oscillations.

[ht]

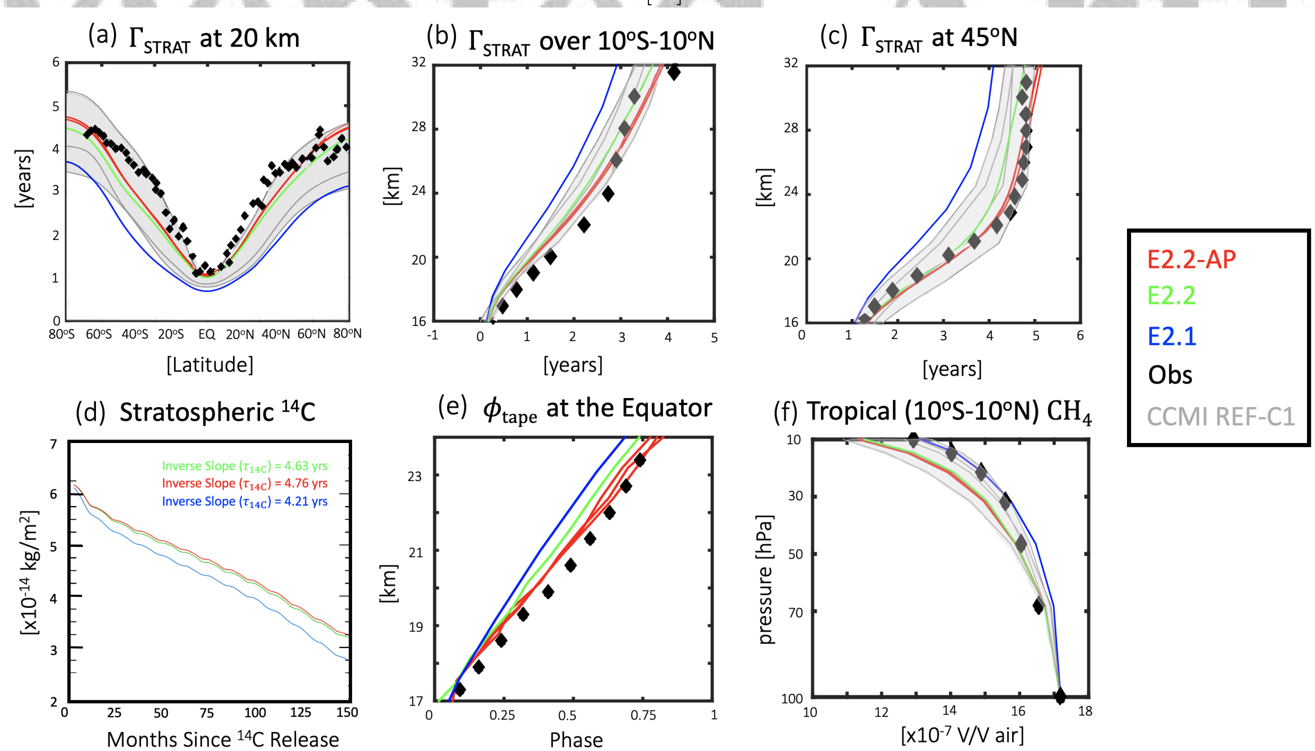

Figure 3. (a) Meridional profile of the stratospheric mean age of air (Γ_{STRAT}) at 20 km. Vertical profiles of Γ_{STRAT}, averaged over 10°S-10°N (b) and evaluated at 45°N (c). The evolution of the (natural log of the) stratospheric mass of ^{14}C is shown in (d), along with the associated stratospheric residence time, inferred as the (inverse) slope after 100 months. Vertical profiles over the tropics (10°S-10°N) of the "tape recorder" phase lag, as inferred from a tape recorder tracer (χ_{tape}) (e) and methane (CH_4) (f), normalized to values at the lowest HALOE level, are also shown. Color lines in all panels denote three ensemble members of E2.2-AP (red), one member of E2.2 (green), and two members of E2.1 (blue) while grey shading denotes the range spanned by the CCMI models (grey lines denote individual models). In-situ based estimates are shown in the black diamonds for Γ_{STRAT} (a-c), χ_{tape} (e) and CH_4 (f).

[ht]

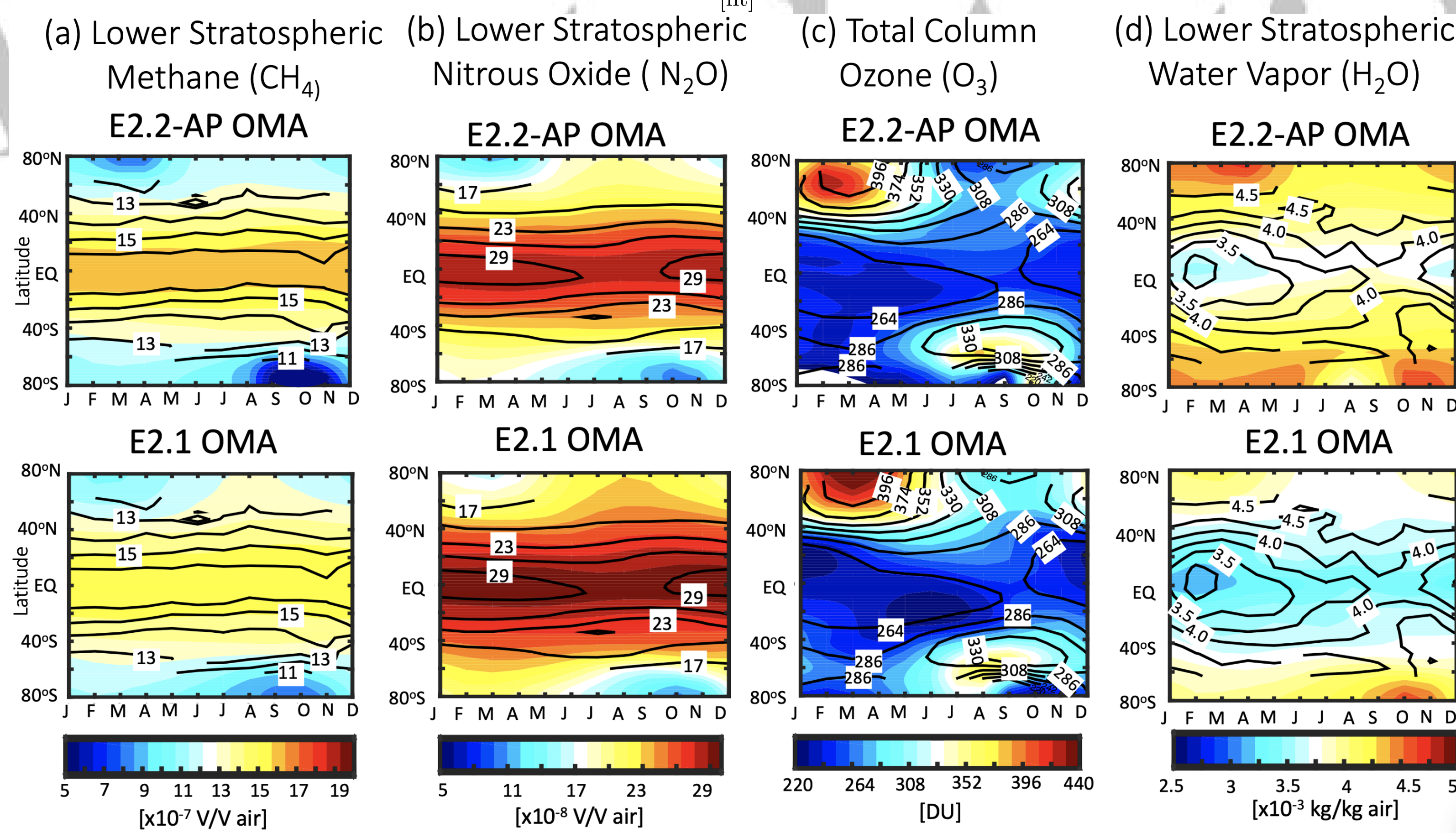

Figure 4. Seasonal cycle of zonally averaged methane (a) (CH_4), (b) nitrous oxide (N_2O), (c) ozone (O_3) and (d) water vapor (H_2O), further averaged over the middle-to-lower stratosphere (30-100 hPa). For the case of ozone (c) the total column is shown. Color contours shows the three-member mean of the E2.2-AP OMA Historical ensemble (top) and one member of the E2.1 OMA Historical simulation (bottom). Black contours denote observed values from HALOE (a,d), AURA-MLS (b) and TOMS/OMI (c). Climatologies have been performed over years 1991-2002 when comparing to HALOE (a,d), over 2005-2014 when comparing to MLS (b) and over 2000-2010 when comparing to TOMS/OMI (c).

[ht]

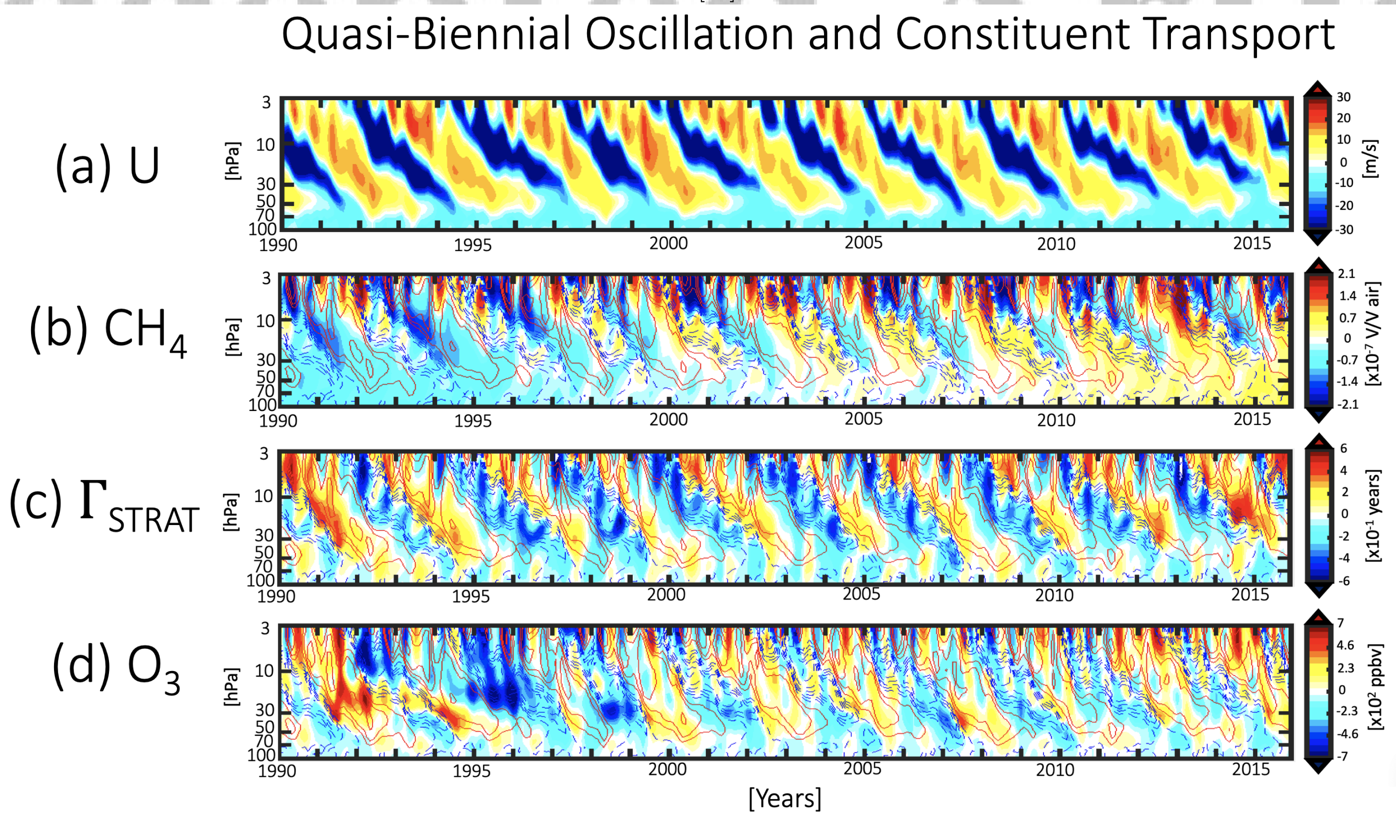

Figure 5. Dynamical and transport signatures of the Quasi-Biennial Oscillation in E2.2-AP as represented in terms of the near-equatorial (5°S-5°N) zonally averaged zonal winds (a) and in terms of climatological anomalies (relative to the 1990-2015 mean) of methane (CH_4) (b), the stratospheric mean age of air (Γ_{STRAT}) (c) and ozone (O_3) (d). Only one ensemble member from the OMA E2.2-AP ensemble is shown so as not to average across different QBO phases among ensemble members. Red and blue contours in (b-d), spaced every 6 m/s, denote negative and positive zonal wind anomalies, respectively.

[ht]

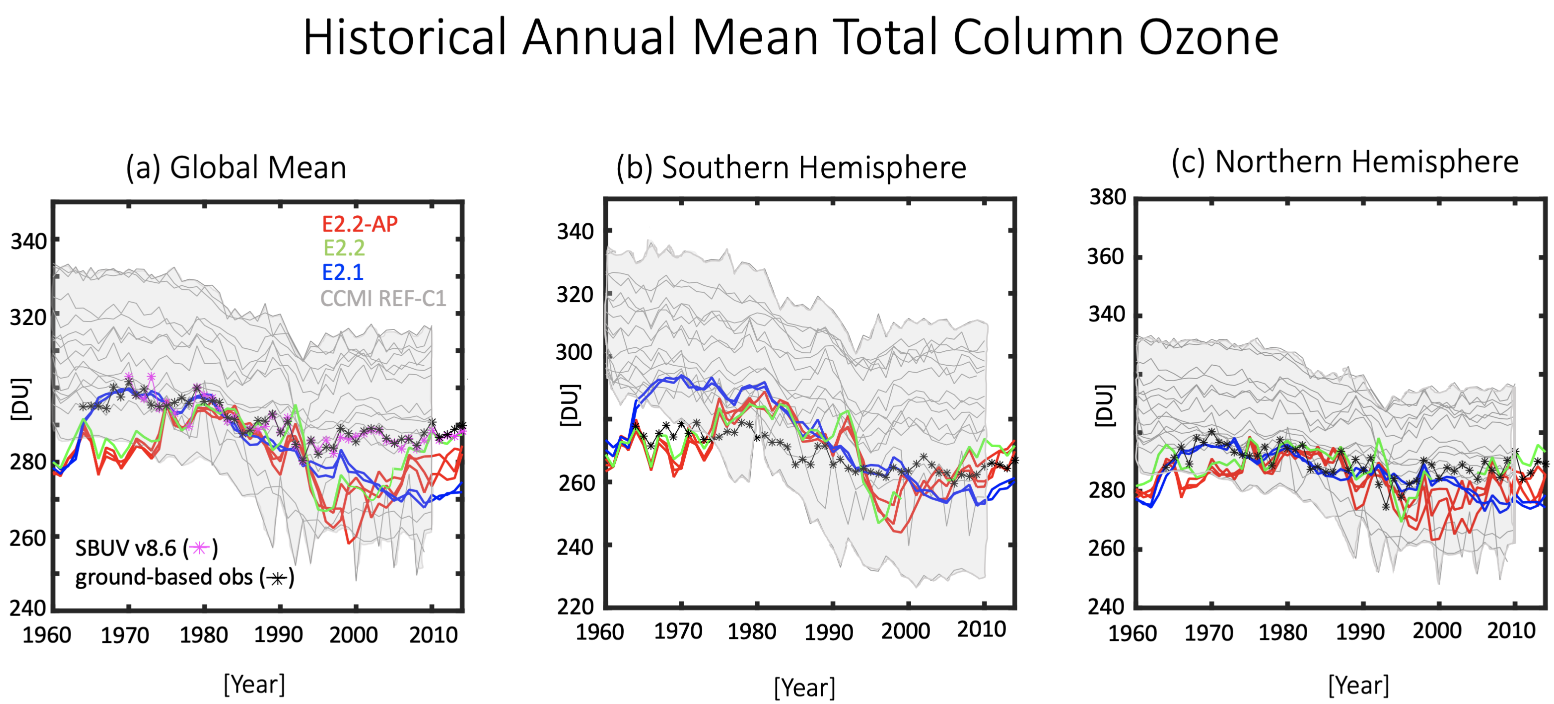

Figure 6. Total column ozone, globally averaged (90°S-90°N) (a) as well as averaged over the Southern Hemisphere (b) and over the Northern Hemisphere (c). Red, green and blue lines denote three ensemble members of E2.2-AP, one member of E2.2, and two members of E2.1, respectively. Grey shading denotes the range spanned by the CCMI models up through 2010, the final year of the REF-C1 experiment (grey lines denote individual models). Ground-based and the SBUV v8.6 merged satellite measurements of global mean total column ozone are shown in (a) in black and magenta, respectively.

[ht]

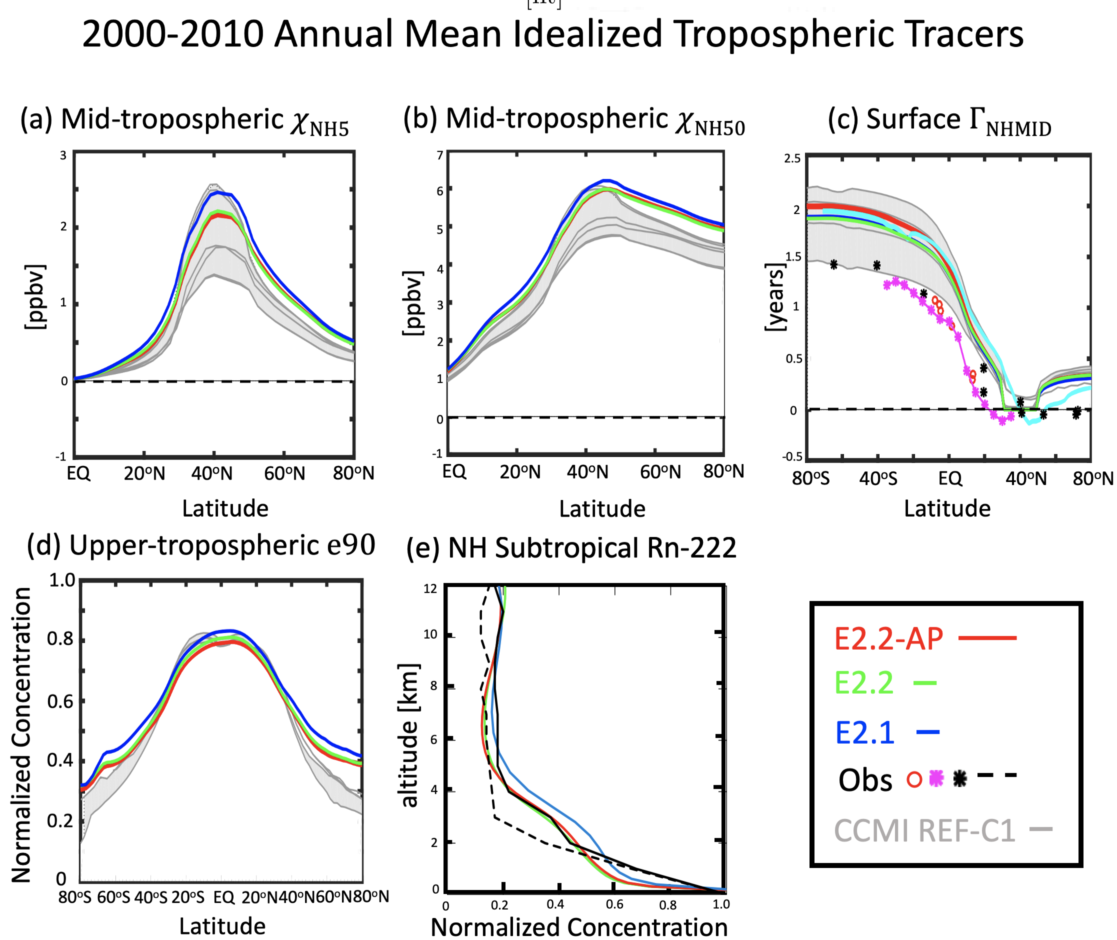

Figure 7. 2000-2010 annual mean climatological meridional profiles of the NH midlatitude surface loss tracers χ_5 and χ_{50} at 500 hPa (a,b), the NH midlatitude mean age, Γ_{NHMID}, evaluated at 900 hPa (c), e90, averaged over 100-500 hPa, and normalized by its surface layer value (d), and northern subtropical profiles of the radon tracer (Rn-222), also normalized by values at the surface (e). Red, green and blue lines denote three ensemble members of E2.2-AP, one member of E2.2, and two members of E2.1, respectively, while grey shading shows the range spanned by the CCMI models (grey lines denote individual models). (Note that Rn-222 was not available for the CCMI models.) The cyan line in (c) denotes the SF_6 age in E2.2-AP, which can be most directly compared to the observed ages (black, magenta, red symbols). For both the models and observations the SF_6 age (a_{SF6}) has been calculated as the time lag satisfying $\chi(\mathbf{r}, t) = \chi_0(t - a_{SF6})$.

[ht]

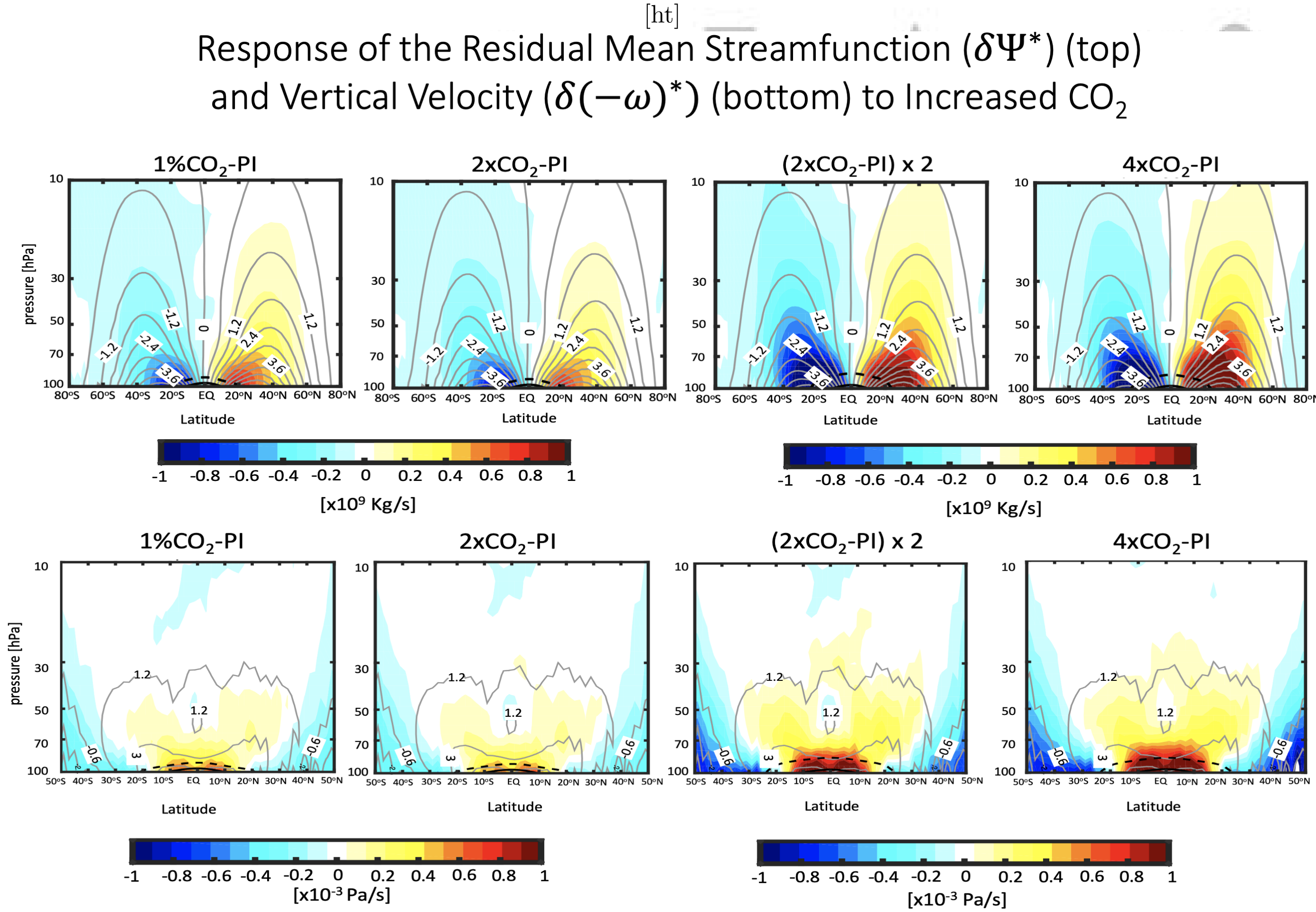

Figure 8. The simulated response in E2.2-AP of the residual mean streamfunction (Ψ^*) (top) and vertical velocity (-ω^*) in the stratosphere (bottom) in response to increased CO_2. Color contours show changes relative to the pre-industrial (PI) control for the 1%CO_2, 2×CO_2 and 4×CO_2 simulations in the first, second and fourth columns, respectively. Grey contours show climatological values for the PI control. Comparisons of double the responses to 2×CO_2 (third column) with the panels in the fourth column indicate the degree to which the circulation responses are linear with CO_2 forcing. (Note that changes in upwelling are depicted in terms of (-ω^*) in the bottom panels in order to reinforce the sense that the circulation is accelerating in response to increased CO_2.)

[ht]

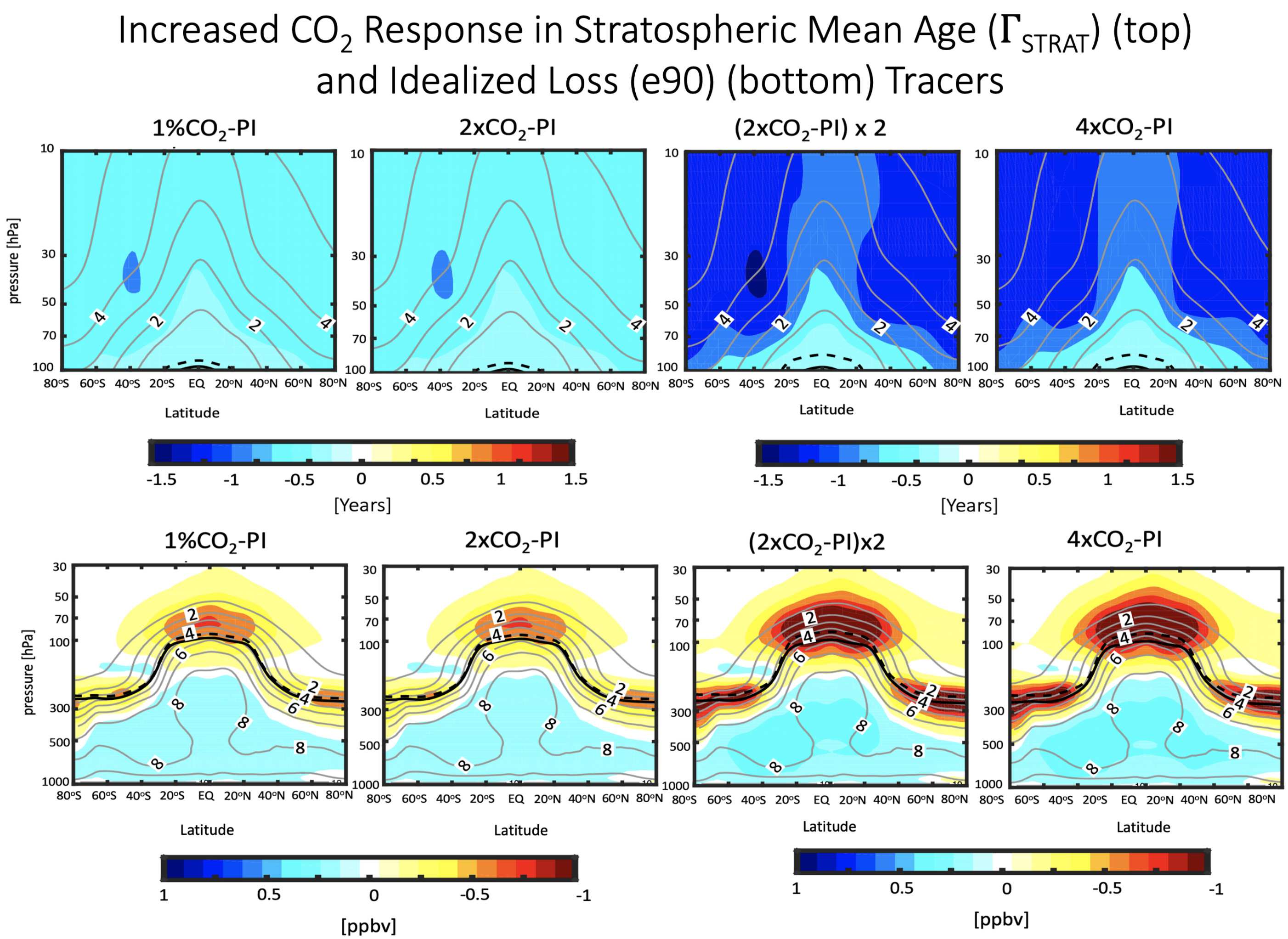

Figure 9. As in Figure 8, except now showing changes in the stratospheric mean age (Γ_{STRAT}) (top) and e90 (bottom). The black dashed and solid thick lines show the climatological annually averaged tropopause for the increased CO_2 and PI control experiments, respectively. Grey contours show climatological values for the PI control.

[ht]

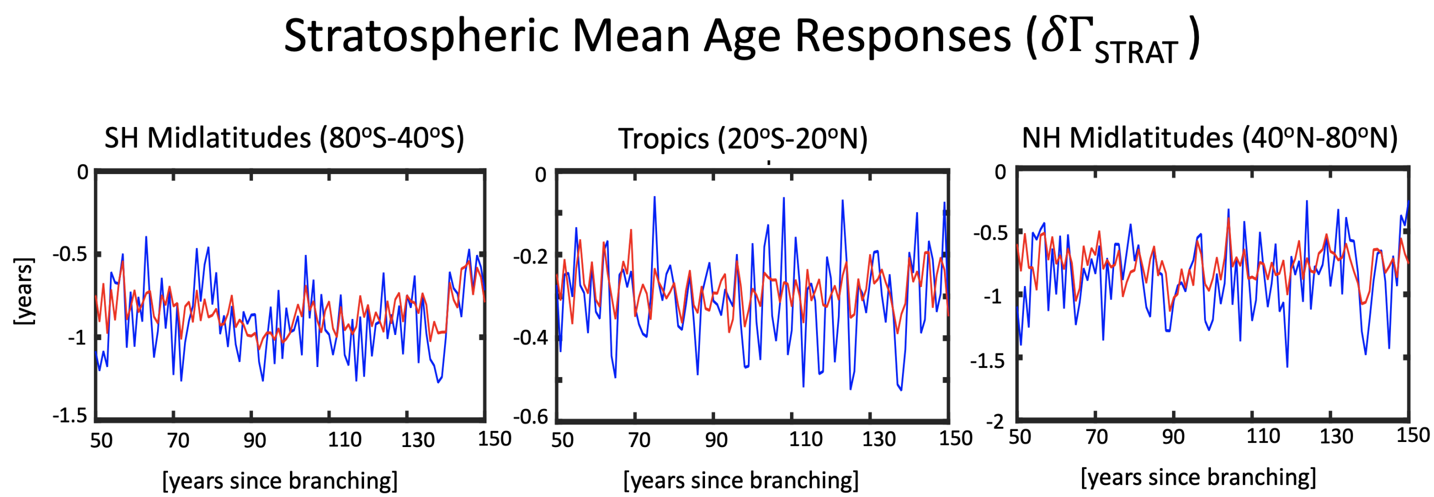

Figure 10. Evolution of changes in the stratospheric mean age of air, Γ_{STRAT}, relative to the PI control for (twice the) response to $2xCO_2$ (blue lines) and $4xCO_2$ (red lines). Results are shown for E2.2-AP for years 50-150 after branching from the underlying PI control. Latitudinal averages at 70 hPa have been performed over SH midlatitudes (80°S–40°S) (left), the tropics (20°S–20°N) (middle), and NH midlatitudes (40°N–80°N) (right).

[ht]

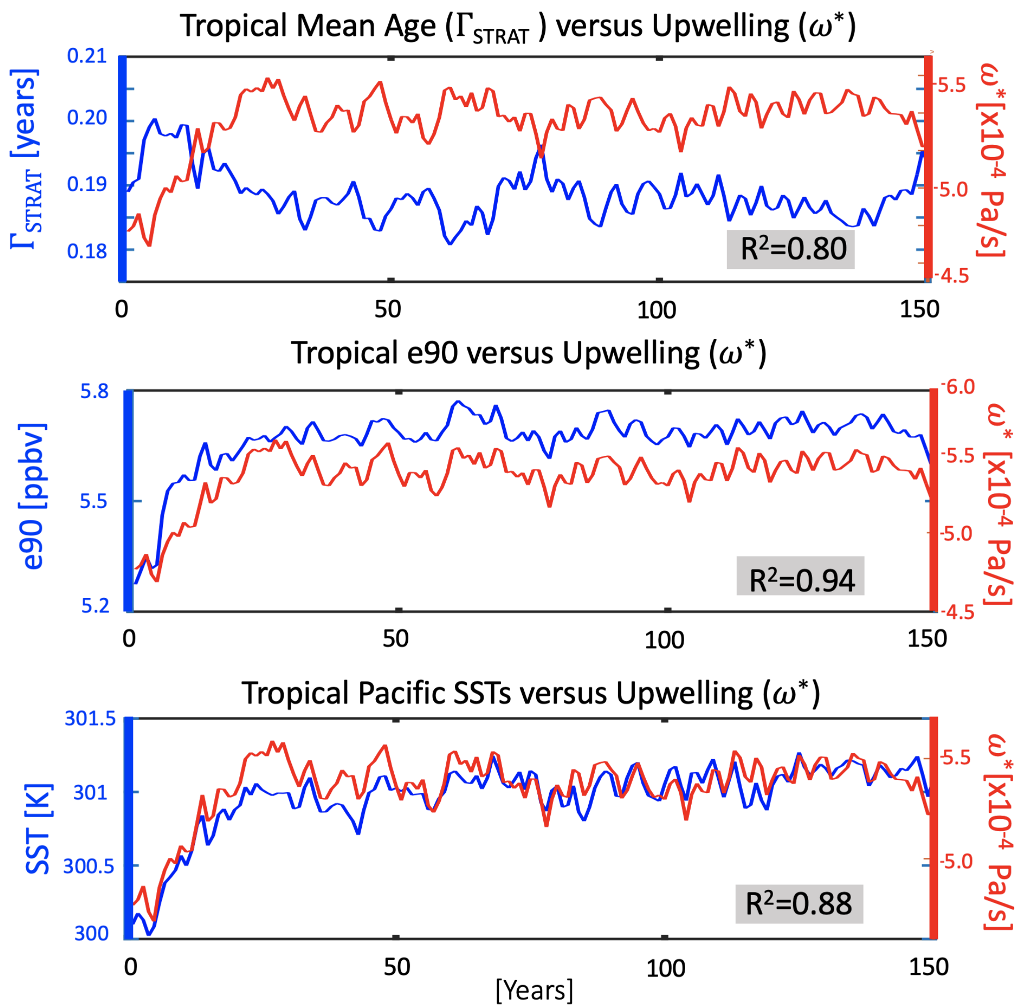

Figure 11. Evolution of lower stratospheric tropical upwelling in E2.2-AP (red lines) compared against the evolution of the stratospheric mean age (Γ_{STRAT}) (blue line, top) and e90 (blue line, middle), also evaluated in the lower stratosphere (i.e. latitudes spanning 10°S-10°N at 70 hPa). The bottom panel compares the evolution of upwelling with tropical Pacific sea surface temperatures, averaged over 10°S-10°N. Correlations of the fields in each panel are shown in the bottom right-hand panels. Note that a four-year moving average has been applied to all timeseries.

[ht]

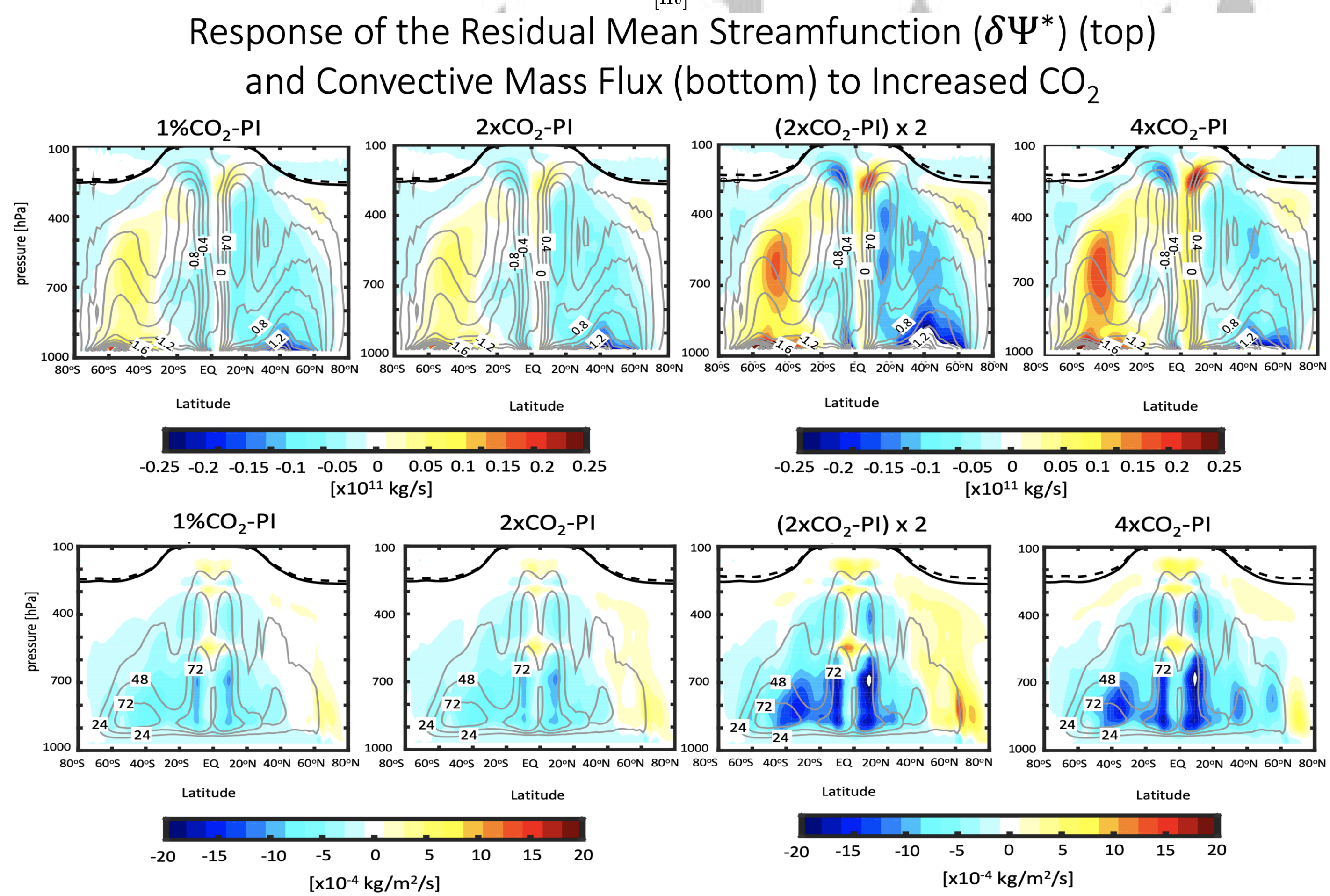

Figure 12. As in Figures 8 and 9, except showing annual mean changes in the residual mean circulation (Ψ^*) evaluated in the troposphere (top) and vertical mass exchange due to (parameterized) convection (bottom), respectively. The black dashed and solid thick lines show the climatological annually averaged tropopause for the increased CO_2 and PI control experiments, respectively. As in previous figures results are shown for E2.2-AP and grey contours show climatological values for the PI control.

[ht]

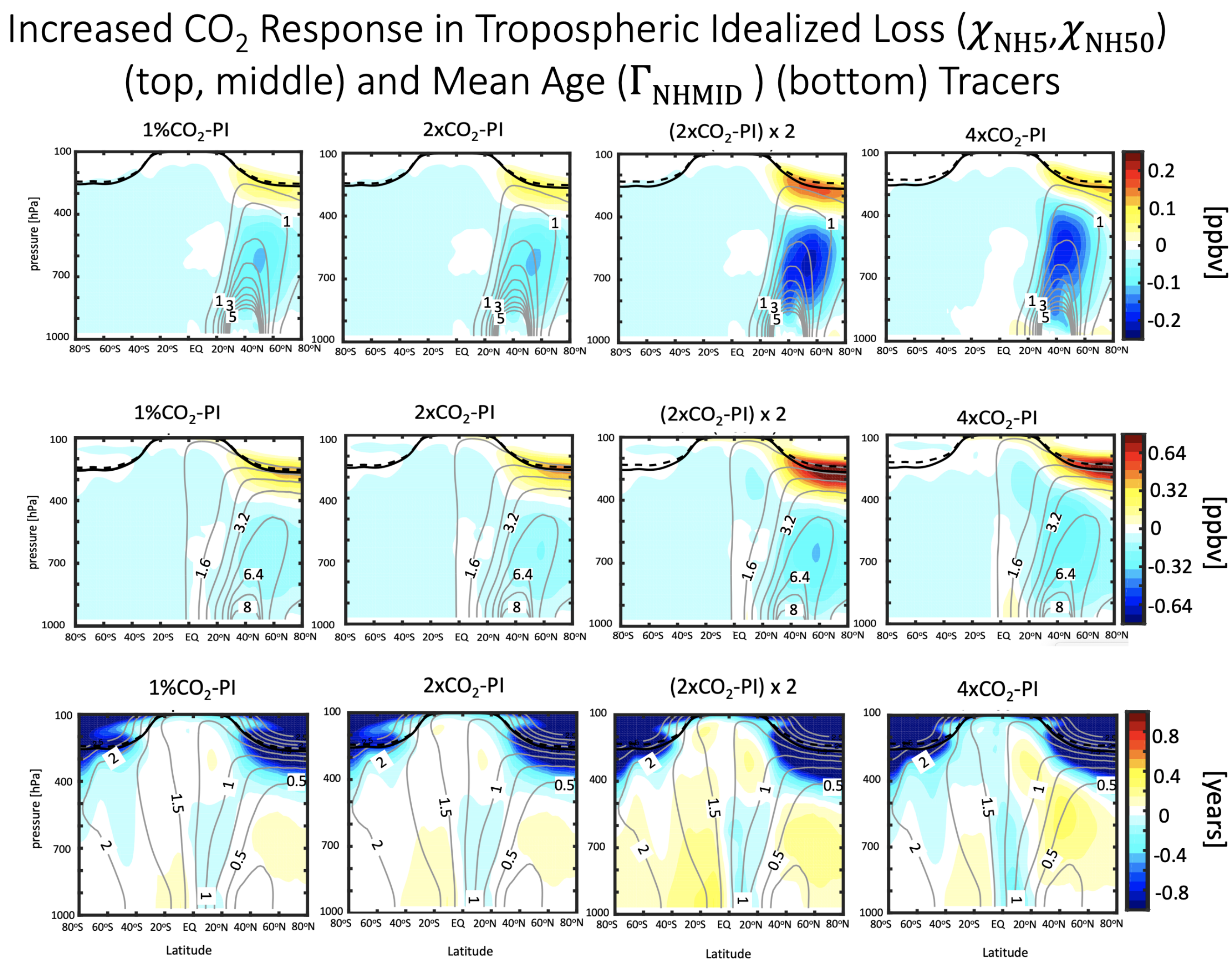

Figure 13. As in Figures 8, 9 and 12, except showing annual mean changes in the 5-day and 50-day idealized loss tracers, $\chi_{NH,5}$ (top) and $\chi_{NH,50}$ (middle), respectively. Changes in the mean transit time since air was last at the NH midlatitude surface (Γ_{NHMID}) are shown in the bottom panels. The black dashed and solid thick lines show the climatological annually averaged tropopause for the increased CO_2 and PI control experiments, respectively. As in previous figures results are shown for E2.2-AP and grey contours show climatological values for the PI control.

[ht]

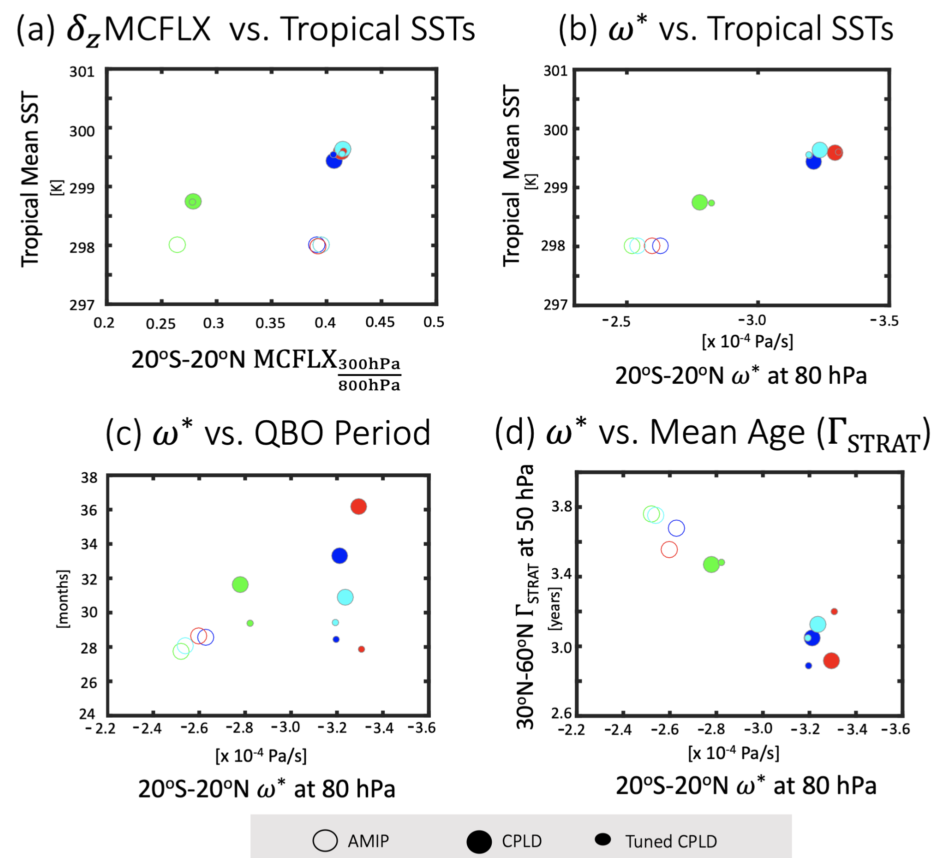

Figure 14. The relationship between climatological mean sea surface temperatures in the tropics and the 300/800 hPa ratio of the vertical mass exchange due to (parameterized) convection (a), tropical SSTs and lower stratospheric (LS) upwelling (ω^* at 80 hPa averaged over 20°S-20°N) (b), LS upwelling and QBO period (c) and LS upwelling versus the NH midlatitude (stratospheric) mean age of air (Γ_{STRAT}) at 50 hPa (d). Circles show various sensitivity experiments using E2.2 (Table 1, rows 10-11). Open (closed) circles denote AMIP (coupled atmosphere-ocean) simulations while green, red, blue and cyan correspond to individual experiment pairs. Large (small) filled circles denote the coupled runs prior to (after) tuning the QBO period.

References

Abalos, M., Orbe, C., Kinnison, D. E., Plummer, D., Oman, L. D., Jöckel, P., ... others (2019). Future trends in stratosphere-to-troposphere transport in CCMI models. *Atmospheric Chemistry and Physics (ACP)*. doi: 10.5194/acp-2019-581

Abalos, M., Randel, W. J., Kinnison, D. E., & Garcia, R. R. (2017). Using the artificial tracer e90 to examine present and future UTLS tracer transport in WACCM. *Journal of the Atmospheric Sciences*, *74*(10), 3383–3403.

Andrews, A. E., Boering, K. A., Daube, B. C., Wofsy, S. C., Loewenstein, M., Jost, H., ... others (2001). Mean ages of stratospheric air derived from in situ observations of CO_2, CH_4, and N_2O. *Journal of Geophysical Research: Atmospheres*, *106*(D23), 32295–32314.

Andrews, D. G., Leovy, C. B., & Holton, J. R. (1987). *Middle Atmosphere Dynamics*. Academic Press.

Appenzeller, C., Holton, J. R., & Rosenlof, K. H. (1996). Seasonal variation of mass transport across the tropopause. *Journal of Geophysical Research: Atmospheres*, *101*(D10), 15071–15078.

Ayarzagüena, B., Charlton-Perez, A. J., Butler, A. H., Hitchcock, P., Simpson, I. R., Polvani, L. M., ... others (2020). Uncertainty in the response of sudden stratospheric warmings and stratosphere-troposphere coupling to quadrupled CO_2 concentrations in CMIP6 models. *Journal of Geophysical Research: Atmospheres*, e2019JD032345.

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

Bushell, A. C., Anstey, J. A., Butchart, N., Kawatani, Y., Osprey, S. M., Richter, J. H., ... Yukimoto, S. (2020). Evaluation of the Quasi-Biennial Oscillation in global climate models for the SPARC QBO-initiative. *Quarterly Journal of the Royal Meteorological Society*, 1–31. doi: 10.1002/qj.3765

Butchart, N., Cionni, I., Eyring, V., Shepherd, T. G., Waugh, D. W., Akiyoshi, H., ... others (2010). Chemistry–climate model simulations of twenty-first century stratospheric climate and circulation changes. *Journal of Climate*, *23*(20), 5349–5374.

Butchart, N., & Scaife, A. A. (2001). Removal of chlorofluorocarbons by increased mass exchange between the stratosphere and troposphere in a changing climate. *Nature*, *410*(6830), 799–802.

CCMVal, S. (2010). *SPARC report on the evaluation of chemistry-climate models, Edited by: Eyring, V., Shepherd, T.G., and Waugh, D.W.* (Tech. Rep.). SPARC report.

Chemke, R., & Polvani, L. M. (2019). Exploiting the abrupt 4×CO_2 scenario to elucidate tropical expansion mechanisms. *Journal of Climate*, *32*(3), 859–875.

Chiodo, G., & Polvani, L. M. (2017). Reduced Southern Hemispheric circulation response to quadrupled CO_2 due to stratospheric ozone feedback. *Geophysical Research Letters*, *44*(1), 465–474.

Chipperfield, M. P., Dhomse, S., Hossaini, R., Feng, W., Santee, M. L., Weber, M., ... Coldewey-Egbers, M. (2018). On the cause of recent variations in lower stratospheric ozone. *Geophysical Research Letters*, *45*(11), 5718–5726.

Danielsen, E. F. (1993). In situ evidence of rapid, vertical, irreversible transport of lower tropospheric air into the lower tropical stratosphere by convective cloud turrets and by larger-scale upwelling in tropical cyclones. *Journal of Geophysical Research: Atmospheres*, *98*(D5), 8665–8681.

Del Genio, A. D., Yao, M.-S., & Jonas, J. (2007). Will moist convection be stronger in a warmer climate? *Geophysical Research Letters*, *34*(16). doi: 10.1029/2007GL030525

Doherty, R. M., Orbe, C., Zeng, G., Plummer, D. A., Prather, M. J., Wild, O., ... Mackenzie, I. A. (2017). Multi-model impacts of climate change on pollution transport from global emission source regions. *Atmospheric Chemistry and Physics*, *17*(23), 14219–14237.

Eichinger, R., Garny, H., Šácha, P., Danker, J., Dietmüller, S., & Oberländer-Hayn, S. (2020). Effects of missing gravity waves on stratospheric dynamics; Part 1: Climatology. *Climate Dynamics*, *54*(5), 3165–3183.

Engel, A., Möbius, T., Bönisch, H., Schmidt, U., Heinz, R., Levin, I., ... others (2009). Age of stratospheric air unchanged within uncertainties over the past 30 years. *Nature Geoscience*, *2*(1), 28–31.

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., & Taylor, K. E. (2016). Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development*, *9*, 1937–1958. doi: 10.5194/gmd-9-1937-2016

Eyring, V., Butchart, N., Waugh, D. W., Akiyoshi, H., Austin, J., Bekki, S., ... others (2006). Assessment of temperature, trace species, and ozone in chemistry-climate model simulations of the recent past. *Journal of Geophysical Research: Atmospheres*, *111*. doi: 10.1029/2006JD007327

Eyring, V., Lamarque, J.-F., Hess, P., Arfeuille, F., Bowman, K., Chipperfiel, M. P., ... others (2013). Overview of IGAC/SPARC Chemistry-Climate Model Initiative (CCMI) community simulations in support of upcoming ozone and climate assessments. *SPARC Newsletter*, *40*(January), 48–66.

Fang, Y., Fiore, A. M., Horowitz, L. W., Gnanadesikan, A., Held, I., Chen, G., ... Levy, H. (2011). The impacts of changing transport and precipitation on pollutant distributions in a future climate. *Journal of Geophysical Research: Atmospheres*, *116*(D18).

Fioletov, V. E., Bodeker, G. E., Miller, A. J., McPeters, R. D., & Stolarski, R. (2002). Global and zonal total ozone variations estimated from ground-based and satellite measurements: 1964–2000. *Journal of Geophysical Research: Atmospheres*, *107*(D22). doi: 10.1029/2001JD001350

Garcia, R. R., & Randel, W. J. (2008). Acceleration of the Brewer-Dobson circulation due to increases in greenhouse gases. *Journal of the Atmospheric Sciences*, *65*(8), 2731–2739.

Geller, L. S., Elkins, J. W., Lobert, J. M., Clarke, A. D., Hurst, D. F., Butler, J. H., & Myers, R. C. (1997). Tropospheric SF_6: Observed latitudinal distribution and trends, derived emissions and interhemispheric exchange time. *Geophysical Research Letters*, *24*(6), 675–678.

Gerber, E. P., Butler, A., Calvo, N., Charlton-Perez, A., Giorgetta, M., Manzini, E., ... others (2012). Assessing and understanding the impact of stratospheric dynamics and variability on the earth system. *Bulletin of the American Meteorological Society*, *93*(6), 845–859.

Gettelman, A., & Sobel, A. H. (2000). Direct diagnoses of stratosphere–troposphere exchange. *Journal of the Atmospheric Sciences*, *57*(1), 3–16.

Grise, K. M., & Polvani, L. M. (2016). Is climate sensitivity related to dynamical sensitivity? *Journal of Geophysical Research: Atmospheres*, *121*(10), 5159–5176.

Grooß, J.-U., & Russell, J. M. (2005, October). Technical note: A stratospheric climatology for O_3, H_2O, CH_4, NO_x, HCl and HF derived from HALOE measurements. *Atmospheric Chemistry and Physics*, *5*(10), 2797–2807. doi: 10.5194/acp-5-2797-2005

Hall, T. M., & Plumb, R. A. (1994). Age as a diagnostic of stratospheric transport. *Journal of Geophysical Research: Atmospheres*, *99*(D1), 1059–1070.

Hall, T. M., & Waugh, D. W. (2000). Stratospheric residence time and its relationship to mean age. *Journal of Geophysical Research: Atmospheres*, *105*(D5), 6773–6782.

Hall, T. M., Waugh, D. W., Boering, K. A., & Plumb, R. A. (1999). Evaluation of transport in stratospheric models. *Journal of Geophysical Research: Atmospheres*, *104*(D15), 18815–18839.

Hardiman, S. C., Butchart, N., & Calvo, N. (2014). The morphology of the Brewer-Dobson circulation and its response to climate change in CMIP5 simulations. *Quarterly Journal of the Royal Meteorological Society*, *140*(683), 1958–1965.

Held, I. M., & Soden, B. J. (2006). Robust responses of the hydrological cycle to global warming. *Journal of climate*, *19*(21), 5686–5699.

Hess, P. G. (2005). A comparison of two paradigms: The relative global roles of moist convective versus nonconvective transport. *Journal of Geophysical Research: Atmospheres*, *110*(D20). doi: 10.1029/2004JD005456

Holton, J. R., & Gettelman, A. (2001). Horizontal transport and the dehydration of the stratosphere. *Geophysical Research Letters*, *28*(14), 2799–2802.

Holzer, M., & Boer, G. J. (2001). Simulated changes in atmospheric transport climate. *Journal of Climate*, *14*(23), 4398–4420.

Holzer, M., & Hall, T. M. (2000). Transit-time and tracer-age distributions in geophysical flows. *Journal of the Atmospheric Sciences*, *57*(21), 3539–3558.

Holzer, M., Orbe, C., & Primeau, F. W. (2012). Stratospheric mean residence time and mean age on the tropopause: Connections and implications for observational constraints. *Journal of Geophysical Research: Atmospheres*, *117*(D12).

Johnston, H. (1989). Evaluation of excess carbon 14 and strontium 90 data for suitability to test two-dimensional stratospheric models. *Journal of Geophysical Research: Atmospheres*, *94*(D15), 18485–18493.

Kelley, M., Schmidt, G. A., Nazarenko, L. S., Bauer, S. E., Ruedy, R., Russell, G. L., ... others (2019). Giss-e2. 1: Configurations and climatology. *Journal of Advances in Modeling Earth Systems*, e2019MS002025.

Kim, D., Del Genio, A. D., & Yao, M.-S. (2013). Moist convection scheme in Model E2. *arXiv preprint arXiv:1312.7496*.

Klonecki, A., Hess, P., Emmons, L., Smith, L., Orlando, J., & Blake, D. (2003). Seasonal changes in the transport of pollutants into the arctic troposphere-model study. *Journal of Geophysical Research: Atmospheres*, *108*(D4).

Lee, Y., Lamarque, J., Flanner, M., Jiao, C., Shindell, D., Berntsen, T., ... others (2013). Evaluation of preindustrial to present-day black carbon and its albedo forcing from Atmospheric Chemistry and Climate Model Intercomparison Project (ACCMIP). *Atmospheric Chemistry and Physics*, *13*(5), 2607–2634.

Levin, I., & Hesshaimer, V. (1996). Refining of atmospheric transport model entries by the globally observed passive tracer distributions of ^{85}krypton and sulfur hexafluoride (SF_6). *Journal of Geophysical Research: Atmospheres*, *101*(D11), 16745–16755.

Li, F., Austin, J., & Wilson, J. (2008). The strength of the Brewer-Dobson circulation in a changing climate: Coupled chemistry–climate model simulations. *Journal of Climate*, *21*(1), 40–57.

Li, F., Stolarski, R. S., Pawson, S., Newman, P. A., & Waugh, D. (2010). Narrowing of the upwelling branch of the Brewer-Dobson circulation and Hadley Cell in chemistry-climate model simulations of the 21st century. *Geophysical Research Letters*, *37*(13).

Lindzen, R. (1984). Gravity waves in the mesosphere. *Dynamics of the Middle Atmosphere*, *3*, 18.

Lorenz, D. J., & DeWeaver, E. T. (2007). Tropopause height and zonal wind response to global warming in the IPCC scenario integrations. *Journal of Geophysical Research: Atmospheres*, *112*(D10).

Lu, J., Chen, G., & Frierson, D. M. W. (2008). Response of the zonal mean atmospheric circulation to El Niño versus global warming. *Journal of Climate*, *21*(22), 5835–5851.

Marvel, K., Schmidt, G. A., Shindell, D., Bonfils, C., LeGrande, A. N., Nazarenko, L., & Tsigaridis, K. (2015). Do responses to different anthropogenic forcings add linearly in climate models? *Environmental Research Letters*, *10*(10), 104010.

McPeters, R., Kroon, M., Labow, G., Brinksma, E., Balis, D., Petropavlovskikh, I., ... Levelt, P. (2008). Validation of the Aura ozone monitoring instrument total column ozone product. *Journal of Geophysical Research: Atmospheres*, *113*(D15).

McPeters, R. D., Bhartia, P., Haffner, D., Labow, G. J., & Flynn, L. (2013). The version 8.6 sbuv ozone data record: An overview. *Journal of Geophysical Research: Atmospheres*, *118*(14), 8032–8039.

Menzel, M. E., Waugh, D., & Grise, K. (2019). Disconnect between Hadley cell and subtropical jet variability and response to increased CO_2. *Geophysical Research Letters*, *46*(12), 7045–7053.

Mote, P. W., Rosenlof, K. H., McIntyre, M. E., Carr, E. S., Gille, J. C., Holton, J. R., ... Waters, J. W. (1996). An atmospheric tape recorder: The imprint of tropical tropopause temperatures on stratospheric water vapor. *Journal of Geophysical Research: Atmospheres*, *101*(D2), 3989–4006.

Murray, L. T., Mickley, L. J., Kaplan, J. O., Sofen, E. D., Pfeiffer, M., & Alexander, B. (2014). Factors controlling variability in the oxidative capacity of the troposphere since the last glacial maximum. *Atmospheric Chemistry and Physics*.

Neu, J. L., & Plumb, R. A. (1999). Age of air in a leaky pipe model of stratospheric transport. *Journal of Geophysical Research: Atmospheres*, *104*(D16), 19243–19255.

Oberländer, S., Langematz, U., & Meul, S. (2013). Unraveling impact factors for future changes in the Brewer-Dobson circulation. *Journal of Geophysical Research: Atmospheres*, *118*(18), 10–296.

O'Gorman, P. A., & Schneider, T. (2008). The hydrological cycle over a wide range of climates simulated with an idealized gcm. *Journal of Climate*, *21*(15), 3815–3832.

Oman, L., Waugh, D. W., Pawson, S., Stolarski, R. S., & Newman, P. A. (2009). On the influence of anthropogenic forcings on changes in the stratospheric mean age. *Journal of Geophysical Research: Atmospheres*, *114*(D3).

Orbe, C., Holzer, M., & Polvani, L. M. (2012). Flux distributions as robust diagnostics of stratosphere-troposphere exchange. *Journal of Geophysical Research: Atmospheres*, *117*(D1).

Orbe, C., Holzer, M., Polvani, L. M., & Waugh, D. (2013). Air-mass origin as a diagnostic of tropospheric transport. *Journal of Geophysical Research: Atmospheres*, *118*(3), 1459–1470.

Orbe, C., Newman, P. A., Waugh, D. W., Holzer, M., Oman, L. D., Li, F., & Polvani, L. M. (2015). Air-mass origin in the Arctic. Part II: Response to increases in greenhouse gases. *Journal of Climate*, *28*(23), 9105–9120.

Orbe, C., Oman, L. D., Strahan, S. E., Waugh, D. W., Pawson, S., Takacs, L. L., & Molod, A. M. (2017). Large-scale atmospheric transport in GEOS replay simulations. *Journal of Advances in Modeling Earth Systems*, *9*(7), 2545–2560.

Orbe, C., Roekel, L. V., Adames, A. F., Dezfuli, A., Fasullo, J., Gleckler, P. J., ... Schmidt, G. A. (2020). Current Representation of Modes of Variability in 6 U.S. Climate Models. *Journal of Climate*. (Submitted)

Orbe, C., Yang, H., Waugh, D., Zeng, G., Morgenstern, O., Kinnison, D., ... others (2019). Large-scale tropospheric transport in the Chemistry-Climate Model Initiative (CCMI) simulations. *Atmos. Chem. Phys.*, *18*, 7217–7235.

Plumb, R. A. (2002). Stratospheric transport. *Journal of the Meteorological Society of Japan. Ser. II*, *80*(4B), 793–809.

Polvani, L. M., Abalos, M., Garcia, R., Kinnison, D., & Randel, W. J. (2018). Significant weakening of Brewer-Dobson circulation trends over the 21st century as a consequence of the Montreal Protocol. *Geophysical Research Letters*, *45*(1), 401–409.

Polvani, L. M., Waugh, D. W., Correa, G. J., & Son, S.-W. (2011). Stratospheric ozone depletion: The main driver of twentieth-century atmospheric circulation changes in the Southern Hemisphere. *Journal of Climate*, *24*(3), 795–812.

Prather, M. J., & Remsberg, E. E. (1993). The atmospheric effects of stratospheric aircraft. report of the 1992 models and measurements workshop. volume 3: Special diagnostic studies.

Randel, W. J., Wu, F., Oltmans, S. J., Rosenlof, K., & Nedoluha, G. E. (2004). Interannual changes of stratospheric water vapor and correlations with tropical tropopause temperatures. *Journal of the Atmospheric Sciences*, *61*(17), 2133–2148.

Rani, K. P., Chandrashekara, M., & Paramesh, L. (2014). Vertical profile of radon and its progeny concentrations and its effect on atmospheric electrical conductivity near the surface of the earth. *International Journal of Advanced Scientific and Technical Research*, *4*(4).

Richter, J. H., Butchart, N., Kawatani, Y., Bushell, A. C., Holt, L., Serva, F., ... others (2019). Response of the Quasi-Biennial Oscillation to a warming climate in global climate models. *Quarterly Journal of the Royal Meteorological Society*.

Rind, D., Jonas, J., Balachandran, N. K., Schmidt, G. A., & Lean, J. (2014). The QBO in two GISS global climate models: 1. Generation of the QBO. *Journal of Geophysical Research: Atmospheres*, *119*(14), 8798–8824.

Rind, D., Lerner, J., Perlwitz, J., McLinden, C., & Prather, M. (2002). Sensitivity of tracer transports and stratospheric ozone to sea surface temperature patterns in the doubled CO_2 climate. *Journal of Geophysical Research: Atmospheres*, *107*(D24), ACL–25.

Rind, D., Lerner, J., Shah, K., & Suozzo, R. (1999). Use of on-line tracers as a diagnostic tool in general circulation model development: 2. Transport between the troposphere and stratosphere. *Journal of Geophysical Research: Atmospheres*, *104*(D8), 9151–9167.

Rind, D., Orbe, C., Jonas, J., Nazarenko, L., Zhou, T., Kelley, M., ... Tausnev, N. (2020). Giss model e2.2: A climate model optimized for the middle atmosphere. part 1: Model structure, climatology, variability and climate sensitivity. *J. Geophys. Res. Atmos.*. doi: 10.1029/2019JD032204

Rind, D., Shindell, D., Lonergan, P., & Balachandran, N. K. (1998). Climate change and the middle atmosphere. Part III: The doubled CO_2 climate revisited. *Journal of Climate*, *11*(5), 876–894.

Rind, D., Suozzo, R., Balachandran, N. K., Lacis, A., & Russell, G. (1988). The GISS global climate-middle atmosphere model. Part I: Model structure and climatology. *Journal of the Atmospheric Sciences*, *45*(3), 329–370.

Scaife, A. A., Karpechko, A. Y., Baldwin, M. P., Brookshaw, A., Butler, A. H., Eade, R., ... others (2016). Seasonal winter forecasts and the stratosphere. *Atmospheric Science Letters*, *17*(1), 51–56.

Schoeberl, M., Douglass, A., Newman, P., Lait, L., Lary, D., Waters, J., ... others (2008). Qbo and annual cycle variations in tropical lower stratosphere trace gases from HALOE and Aura MLS observations. *Journal of Geophysical Research: Atmospheres*, *113*(D5).

Seviour, W. J. M., Hardiman, S. C., Gray, L. J., Butchart, N., MacLachlan, C., & Scaife, A. A. (2014). Skillful seasonal prediction of the southern annular mode and Antarctic ozone. *Journal of Climate*, *27*(19), 7462–7474.

Sherwood, S. C., & Dessler, A. E. (2000). On the control of stratospheric humidity. *Geophysical Research Letters*, *27*(16), 2513–2516.

Shindell, D. T., Pechony, O., Voulgarakis, A., Faluvegi, G., Nazarenko, L., Lamarque, J.-F., ... Schmidt, G. A. (2013, March). Interactive ozone and methane chemistry in GISS-e2 historical and future climate simulations. *Atmospheric Chemistry and Physics*, *13*(5), 2653–2689. Retrieved from https://doi.org/10.5194/acp-13-2653-2013 doi: 10.5194/acp-13-2653-2013

Sigmond, M., Siegmund, P. C., Manzini, E., & Kelder, H. (2004). A simulation of the separate climate effects of middle-atmospheric and tropospheric CO_2 doubling. *Journal of climate*, *17*(12), 2352–2367.

Tweedy, O. V., Kramarova, N. A., Strahan, S. E., Newman, P. A., Coy, L., Randel, W. J., … Frith, S. M. (2017). Response of trace gases to the disrupted 2015-2016 Quasi-Biennial Oscillation. *Atmospheric Chemistry & Physics*, *17*(11).

Vecchi, G. A., & Soden, B. J. (2007). Global warming and the weakening of the tropical circulation. *Journal of Climate*, *20*(17), 4316–4340.

Waugh, D. W., Crotwell, A. M., Dlugokencky, E. J., Dutton, G. S., Elkins, J. W., Hall, B. D., … others (2013). Tropospheric SF_6: Age of air from the Northern Hemisphere midlatitude surface. *Journal of Geophysical Research: Atmospheres*, *118*(19), 11–429.

Yang, H., Waugh, D. W., Orbe, C., Patra, P. K., Jöckel, P., Lamarque, J.-F., … Dlugokencky, E. J. (2019). Evaluating simulations of interhemispheric transport: Interhemispheric exchange time versus SF_6 age. *Geophysical Research Letters*, *46*(2), 1113–1120.

Yoo, C., & Son, S.-W. (2016). Modulation of the boreal wintertime Madden-Julian Oscillation by the stratospheric Quasi-Biennial Oscillation. *Geophysical Research Letters*, *43*(3), 1392–1398.