Large-Scale Atmospheric Transport in GEOS Replay Simulations

Clara Orbe [1,2,5], Luke D. Oman [3], Susan E. Strahan [3,4], Darryn W. Waugh [2], Steven Pawson [5], Lawrence L. Takacs [5,6] and Andrea M. Molod [5]

[1]GESTAR
[2]Department of Earth and Planetary Sciences, Johns Hopkins University, Baltimore, Maryland, USA.
[3]Atmospheric Chemistry and Dynamics Laboratory, NASA Goddard Space Flight Center, Greenbelt, Maryland, USA.
[4]Universities Space Research Association
[5]Global Modeling and Assimilation Office, NASA Goddard Space Flight Center, Greenbelt, Maryland, USA.
[6]Science Systems and Applications Inc.

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

Abstract

Offline chemical transport models (CTMs) have traditionally been used to perform studies of atmospheric chemistry in a fixed dynamical environment. An alternative to using CTMs is to constrain the flow in a general circulation model using winds from meteorological analyses. The Goddard Earth Observing System (GEOS) "replay" approach involves reading in analyzed fields every six hours and recomputing the analysis increments, which are applied as a forcing to the meteorology at every model time step. Unlike in CTM, all of the subgrid-scale processes are recalculated on-line so that they are consistent with the large-scale analysis fields, similar in spirit to "nudged" simulations, in which the online meteorology is relaxed to the analysis. Here we compare the transport of idealized tracers in different replay simulations constrained with meteorological fields taken from The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2). We show that there are substantial differences in their large-scale stratospheric transport, depending on whether analysis fields or assimilated fields are used. Replay simulations constrained with the instantaneous analysis fields produce stratospheric mean age values that are up to 30% too young relative to observations; by comparison, simulations constrained with the time-averaged assimilated fields produce more credible stratospheric transport. Our study indicates that care should be taken to correctly configure the model when the replay technique is used to simulate stratospheric composition.

1 Introduction

The chemical and radiative properties of the troposphere and lower stratosphere are strongly influenced by the distributions of greenhouse gases (GHG) and ozone-depleting substances (ODS). Studies on atmospheric composition often use models constrained with analyzed meteorological fields in order to understand the influence of meteorology on the trends and variability of various GHG and ODS. However, transport errors associated with the meteorological fields themselves, as well as the methods by which they are prescribed, can result in unrealistic simulations of both stratospheric and tropospheric composition [e.g., *Schoeberl et al.*, 2003; *Meijer et al.*, 2004; *Prather et al.*, 2008].

Transport errors in computations using data assimilation system (DAS) winds tend to accumulate with time. In the stratosphere the use of assimilated winds can produce a transport circulation that is too fast due to excessive mixing [e.g., *Schoeberl et al.*, 2003; *Legras et al.*, 2004; *Pawson et al.*, 2007]. This is because assimilated winds are characterized by high frequency fluctuations associated with the insertion of data that can induce spurious transport in the lower stratosphere [e.g., *Weaver et al.*, 1993; *Tan et al.*, 2004]. One approach to reducing these errors is to improve the assimilated fields themselves, either by changing the assimilation system or the underlying model [*Monge-Sanz et al.*, 2013].

Transport simulations using assimilated winds are also very sensitive to how the meteorological fields are prescribed. For example, one limitation with using offline chemical transport models (CTMs), is that the timescales of many atmospheric processes are less than the intervals at which the meteorological fields, used to force the model, are archived (typically six hours). Since the lack of resolved sub six-hourly meteorological features can result in large inaccuracies in transport, CTMs currently tend to use more frequently sampled (three-hourly) archived fields [e.g., *Rasch et al.*, 1997].

In general, studies have shown that more frequent temporal sampling and, in some cases, temporal averaging tends to produce smoother, more balanced fields and, correspondingly, more credible stratospheric transport [e.g., *Waugh et al.*, 1997; *Legras et al.*, 2004; **?**; *Pawson et al.*, 2007]. Partly because of these improvements, the performance of offline models has improved substantially in recent years in terms of how well they represent both stratospheric and tropospheric constituents [*Ziemke et al.*, 2014; *Strahan et al.*, 2016; *Strode et al.*, 2016].

While changes in the temporal sampling and averaging of assimilated winds have substantially improved the representation of stratospheric transport in CTM simulations, one major limitation of using CTMs is the accuracy of subgrid-scale transport (e.g. convective mass fluxes and boundary layer mixing), which depend sensitively on the parameterizations and resolution of the model used to produce the assimilated fields. For example, *Yu et al.* [2017] show that offline simulations that are performed at a coarser grid than the parent GCM of the driving meteorological fields are associated with large transport errors related to both the temporal and spatial averaging of the analysis vertical winds. Furthermore, in cases where the convective mass fluxes are taken from the same analysis as the large-scale flow, it is not obvious how they should be rescaled to be consistent with the native grid of the CTM [*Prather et al.*, 2008].

An alternative tool to using CTMs is to constrain the flow in a general circulation model using winds from meteorological analyses. Compared to in a CTM, all of the subgrid-scale processes are recalculated on-line so that they are consistent with both the resolution and convective parameterization of the large-scale analysis fields. One such approach, developed at the NASA Global Modeling Assimilation Office, involves running the Goddard Earth Observing System (GEOS) GCM in "replay" mode, wherein the model reads in fields from a pre-existing analysis (e.g. The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2)) every six hours and recomputes the analysis increments, which are applied as a forcing to the meteorology at every time step over the six hour replay interval.

The replay technique also provides a way to perform constrained meteorology simulations using the most recent version of the model (i.e. updated chemistry schemes and subgrid-scale parameterizations), which can be desirable for studying coupled atmosphere-aerosol and atmosphere-chemistry interactions [*Colarco et al.*, 2010; *Strode et al.*, 2015]. In these respects, therefore, GEOS replay simulations are similar in spirit to so-called "nudged" simulations in which an online GCM is relaxed to the meteorological fields from an external analysis (e.g. the Whole Atmosphere Community Climate Model (WACCM) SD simulations described in *Kunz et al.* [2011] and *Lamarque et al.* [2012]).

Similar to the errors associated with offline simulations, it is possible that GEOS replay simulations may also feature unrealistic stratospheric transport properties related to how the flow is prescribed. In particular, how sensitive are GEOS replay simulations to using instantaneous versus time-averaged prescribed meteorological fields? Furthermore, what are the transport differences between replay simulations constrained with (dynamically balanced) assimilated fields versus (unbalanced) analysis fields? We address these questions by comparing the large-scale stratospheric and tropospheric transport properties among various GEOS Version 5 (GEOS-5) [*Rienecker et al.*, 2008; *Molod et al.*, 2015] replay simulations constrained both with instantaneous analysis fields and with time-averaged assimilated fields, both of which are standard MERRA-2 products.

The goal of this study is to assess the credibility of using the replay technique to simulate atmospheric composition, with a focus on large-scale atmospheric transport on long timescales (i.e. months to years) in the stratosphere and the troposphere, including stratosphere-troposphere-exchange. We diagnose transport using idealized tracers, many of which were included in recent Chemistry Climate Modeling Initiative (CCMI) simulations [*Eyring et al.*, 2013] and that provide tracer-independent diagnostics of the flow independent of chemistry and emissions. These include two tracers that quantify stratospheric transport, three tracers that examine transport from the Northern Hemisphere (NH) midlatitude surface, as well as one tracer that quantifies stratosphere-troposphere exchange. Following a brief exposition of the methodology in Section 2 we present results in Sections 3 and 4.

2 Methods

2.1 Description of Replay

The replay technique was originally developed as a way to assess the impact of different underlying model changes on GEOS data assimilation experiments by providing a computationally efficient way to compare the analysis increments among different runs. The GEOS Replay framework is therefore very similar to the standard GEOS DAS procedure in the sense that it uses the same Incremental Analysis Update (IAU) technique that is used to apply the analysis as a correction to the background state [*Bloom et al.*, 1996]. Recall that, in the context of data assimilation, “analysis" fields refer to the fields resulting from the Gridpoint Statistical Interpolation (GSI) analyses, while the “assimilated" fields refer to the result of applying the IAU. The main difference between the replay framework and the standard GEOS DAS, however, is that it uses a *pre-existing analysis* to produce the IAU increments, thus producing an assimilation that is a blend of the analysis combined with the particular version of GEOS used in the simulation. By construction, a GEOS replay simulation would identically produce MERRA-2 assimilated fields if the same version of the model that produced MERRA-2 was used.

As illustrated in Figure 1 (panel a) a GEOS replay simulation first performs a five-hour forecast starting at 21z during the “predictor" segment of the IAU. An increment δX is then calculated as the difference between a time-averaged background state $\overline{X}_{bkg}|_{0z}$ centered about 0z and the pre-existing analysis field X_{ana} at 0z (e.g. MERRA-2). The model then backtracks to 21z and uniformly applies the increment δX to the background state $X_{bkg}|_{21z}$ over a six hour “corrector" interval. Note that the time-averaged background state $\overline{X}_{bkg}|_{0z}$, from which the increment is calculated, is a linear four-hour time average (synoptic time ± 2 hours) that is used to partly suppress resonance frequencies shorter than the IAU six-hour cycle. The four-hour time averaging window is chosen to suppress the most unstable modes, which have characteristic frequencies two-thirds the IAU replay cycle (or four hours).

The averaging of the background state $\overline{X}_{bkg}$, which is used to calculate the increment δX, is designed to inhibit a wave characteristic of the IAU that results in substantial high frequency variability in the assimilated fields that does not coincide with the analysis times. It is not obvious, however, that smoothing the background state sufficiently damps the oscillations already present in the underlying analysis (that result in spurious vertical transport in the lower stratosphere). Furthermore, the degree of consistency among different variables is weaker in the analysis fields, as these are more strongly drawn towards the observations compared to the assimilated fields, which are generated by the model. Therefore, in this study, we directly compare the difference between calculating δX based on the analysis fields versus the (balanced) assimilated fields. Because the analysis and assimilated fields also differ in terms of their temporal sampling – the analysis fields are instantaneous, while the assimilated fields are time-averaged – we also consider the impact of time-averaging the prescribed meteorological fields on large-scale transport. This is described next in more detail.

2.2 Model Simulations

In total five simulations are performed and constrained with MERRA-2 fields for years 1980-2010 (Table 1). These include three online GEOS-5 replay simulations and two integrations using offline models. In the first replay simulation, hereafter referred to RAna, the increments are calculated using six-hourly instantaneous analysis fields X_{ana}, which are available from MERRA-2 at the synoptic times of 00 GMT, 06 GMT, 12 GMT and 18 GMT. The simulation is carried out as described in the previous section and is depicted schematically in Figure 1a.

The other two replay simulations are performed use three-hourly time-averaged assimilated meteorological fields X_{asm}, which are available from MERRA-2 as time averages centered about 01:30 GMT, 04:30 GMT, 07:30 GMT, and so on. We perform two simulations

using X_{asm}, which differ only in terms of the frequency with which the assimilated fields are used to constrain the simulations. The first simulation (RAs6) performs the replay cycle every six hours. A five hour forecast, initialized at 19:30 GMT, is used to calculate the increment δX as the difference between a time-averaged background state $\overline{X}_{bkg}|_{22:30}$ centered about 22:30 GMT and the pre-existing assimilated field X_{asm} at 22:30 GMT. The model then backtracks back to 19:30 GMT and linearly applies the increment δX to the background state $X_{bkg}|_{22:30}$ over a six hour interval (Figure 1b).

The second replay simulation using the three-hourly time-averaged assimilated fields (RAs3) performs the replay cycle every three hours. Starting from a 2.5-hour forecast initialized at 00 GMT the increment δX is calculated as the difference between the time-averaged background state $\overline{X}_{bkg}|_{01:30}$ centered about 01:30 GMT and the pre-existing assimilated field X_{asm} (also at 01:30 GMT). After backtracking to 00 GMT the increment δX is linearly applied to the background state $X_{bkg}|_{01:30}$ over a three-hour interval (Figure 1c). Note that because the replay cycle in RAs3 is half that in RAs6 (and RAna) a two-hour (± 1 hour) background averaging interval is needed to produce $\overline{X}_{bkg}$.

All three GEOS-5 replay simulations are replayed to MERRA-2 zonal and meridional winds, temperature and surface pressure while all other dynamical variables and physics are recalculated online. The simulations are performed at the same C90 cubed sphere horizontal resolution [*Putnam and Lin*, 2007], corresponding approximately to 1° latitude by 1.25° longitude, and constrained by MERRA-2 fields (Table 1). Note that, because all simulations are performed at a coarser resolution compared to the analysis, the increment is first computed online on the 0.625° x 0.5° analysis grid and then interpolated to the coarser C90 grid of the simulation. Convective transport is recomputed every model time step and at the resolution that is consistent with the approximately 1° by 1.25° resolution at which the simulation is performed. Because the convective mass fluxes are recalculated online they reflect the Relaxed-Arakawa Schubert (RAS) convective scheme in its current implementation in the model [*Moorthi and Suarez*, 1992].

In addition to the GEOS-5 replay simulations, two offline simulations are also performed, both of which are constrained with MERRA-2 assimilated fields. The first one uses the NASA Global Modeling Initiative (GMI) three-dimensional chemical transport model, which has been evaluated by several studies in terms of its large-scale transport properties in the stratosphere [*Strahan et al.*, 2007, 2016] and the troposphere [*Waugh et al.*, 2013]. The integration with the NASA GMI CTM, hereafter referred to as the GMI-CTM simulation, is run at a horizontal resolution (1° latitude by 1.25° longitude) and vertical grid (72 vertical levels spanning the surface to 0.01 hPa) that are comparable with the replay simulations. Consistent with its use in previous studies, it is constrained with three-hourly time-averaged MERRA-2 assimilated fields (χ_{asm}) [*Strahan et al.*, 2016]. The second offline simulation uses the NASA GEOS Chemical Transport Model (GEOS-CTM) [*Kouatchou et al.*, 2015] which uses the exact same dynamical and chemical modules used in the GEOS-5 replay simulations, in addition to independent components for offline convective and diffusive transport. The GEOS-CTM simulation is constrained with three-hourly instantaneous MERRA-2 assimilated fields.

Our focus in this study is on large-scale climatological transport, which we assess in terms of ten-year climatological means (denoted throughout using overbars) over the time period 2000-2010. All quantities are based on monthly-mean output, with the exception of daily tracer output which is used to examine mixing between the subtropical and high latitude stratosphere in Section 3. Statistical significance of the differences among all climatological mean quantities, relative to internal variability, is assessed using $\pm\sigma_{\chi}$, where σ denotes the standard deviation of each seasonal mean tracer (or dynamical field) over the climatological averaging period.

2.3 Idealized Tracers

Several of the idealized tracers examined in this study, shown in Table 2, are identical to the tracers presented in *Orbe et al.* [2016, 2017]. The first two tracers comprise one set that is designed to assess large-scale transport in the stratosphere. We first carry an idealized global "clock" or ideal age tracer Γ_{GLB} [e.g., *Thiele and Sarmiento*, 1990] that is defined with respect to all grid points in the first model level (Table 1, row 2). The ideal age tracer is initially set to a value of zero throughout the troposphere and thereafter held to zero over the entire Earth's surface and subject to a constant aging of 1 year/year throughout the atmosphere. The statistically stationary value of $\Gamma_{\mathrm{GLB}}(\mathbf{r})$, the mean age, is equal to the average time since the air at a region $\mathbf{r}$ in the stratosphere last contacted the Earth's surface, and is hereafter referred to as the stratospheric mean age [*Hall et al.*, 1999].

As in *Hall et al.* [1999] we also compare the propagation of an annually periodic oscillation in tracer mixing ratio from the tropical tropopause into the stratosphere (Table 2, row 3). Specifically, we prescribe a sinusoid in a mixing ratio over 10°S-10°N at 100 mb that has a maximum at October, consistent with the seasonality of water vapor-based estimates of the tape recorder at the tropical tropopause [*Mote et al.*, 1996]. The amplitude A and phase lag τ_ω from the tape recorder tracer, hereafter referred to as χ_{tape}, is then compared among the replay simulations and with observations from *Hall et al.* [1999], as described in the next section. During the course of a multi-year-long simulation the near-tropopause gradients of χ_{tape} weaken substantially, since that tracer is not subject to any stratospheric or tropospheric loss. For that reason, we focus on the evolution of χ_{tape} over the first five years of the simulation (1980-1985). Note that this is unlike all of the other variables, which are represented in terms of their 2000-2010 ten-year climatological means. Furthermore, the distribution of χ_{tape} is not shown for GMI-CTM, as that tracer was not included in the simulation when it was performed.

The second class of tracers is designed to assess large-scale transport in the troposphere, as well as stratosphere-troposphere-exchange (Table 2, rows 4-6). Three of the tracers, $\chi_5(\Omega_{\mathrm{MID}})$, $\chi_{50}(\Omega_{\mathrm{MID}})$, and Γ_{NH}, have zonally uniform boundary conditions defined over the same Northern Hemisphere (NH) surface region over midlatitudes, Ω_{MID}, defined as the first model level spanning all grid points between 30°N and 50°N. The tracers χ_5 and χ_{50}, referred to throughout as the 5-day and 50-day idealized loss tracers (χ_5 and χ_{50}), are fixed to a value of 100 ppbv over Ω_{MID} and undergo spatially uniform exponential loss at rates of 5 days^{-1} and 50 days^{-1}, respectively.

The third tropospheric tracer, Γ_{NH}, is similar in spirit to the stratosphere mean age tracer discussed earlier Γ_{GLB}, except that is defined with respect to the NH midlatitude surface Ω_{MID} [*Waugh et al.*, 2013]. Thus, $\Gamma_{\mathrm{NH}}(\mathbf{r})$, is equal to the average time since the air at a region $\mathbf{r}$ last contacted the NH midlatitude surface, Ω_{MID}, and provides a richer description of transport compared to hemispherically integrated transport quantities like the interhemispheric exchange time [*Levin and Hesshaimer*, 1996; *Geller et al.*, 1997]. Finally, the fifth tracer , χ_{STE}, is set to a constant value of 200 ppbv above 80 mb and undergoes spatially uniform exponential loss at a rate of 25 days^{-1} (Table 2, row 7). Similar tracers examined in previous studies have included stratosphere-to-troposphere air-mass origin tracers [*Orbe et al.*, 2013] and the e90 tracer defined in *Prather et al.* [2011].

2.4 Observations

The stratospheric mean age Γ_{GLB} and tape recorder phase lag τ_ω and amplitude A in the simulations are compared with observations previously reported in *Hall et al.* [1999] and *Engel et al.* [2009]. HALOE water vapor measurements are used to infer the tape recorder amplitude A and phase lag τ_ω as in *Mote et al.* [1996] and are identical to the data presented in Figure 16 in *Hall et al.* [1999]). For more details please see *Hall et al.* [1999].

A combination of observations are used for constraining Γ_{GLB} in the simulations. These include the meridional profile of stratospheric mean age observations based on ER2 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 the stratospheric mean age at midlatitudes are taken from the data presented in *Andrews et al.* [2001] and *Engel et al.* [2009]. Specifically, in the tropics we use an average of Γ_{GLB} inferred both from CO_2 and sulfur hexafluoride (SF_6) in situ measurements sampled between 15.2 km-34.2 km and 15.2 km-34 km, respectively. The vertical profiles of Γ_{GLB} at midlatitudes are inferred only using CO_2 measurements from latitudes spanning 34°N-44°N and altitudes between 11.1 km-35.1 km (also from *Andrews et al.* [2001]; *Engel et al.* [2009]).

3 Large-Scale Stratospheric Transport

3.1 Comparison of Stratospheric Mean Age Γ_{GLB}

We first compare the 2000-2010 zonally and annually averaged stratospheric mean age distributions among the two offline model simulations, the GMI-CTM (Figure 2a) and GEOS-CTM (Figure 2b), and two of the replay simulations, RAna (Figure 2c) and RAs6 (Figure 2d). In all cases, the mean age Γ_{GLB} shows the effect of upwelling in the tropics, captured by the large upward bulge of isopleths at low latitudes, with the lowest ages occurring in the tropical lower stratosphere and the oldest ages aloft, consistent with observations [*Hall et al.*, 1999]. The mean age in all simulations is also characterized by sharp latitudinal gradients in the lower stratosphere between ~10° and ~20° from the equator, consistent with observations of trace gases showing that the tropical lower stratosphere is weakly isolated from midlatitudes [*Mote et al.*, 1996; *Volk et al.*, 1996]. This suggests that the isolation of the tropical pipe is more or less well represented in both the offline and replay simulations.

While the shape of Γ_{GLB} is consistent among the simulations, the stratospheric mean age values in the RAna simulation are younger than the ages in both offline CTM simulations. Comparisons of all the simulations with observationally-derived estimates of the mean age at 20 km reveal that Γ_{GLB} in the RAna simulation is too young by ~ 1 year (or 25%) (Figure 3a) and that the differences at middle and high latitudes are statistically significant relative to internal variability (σ_Γ). By comparison, Γ_{GLB} in the GMI-CTM, GEOS-CTM, RAs3 and RAs6 simulations compare well with the observations.

Among the simulations constrained with the assimilated fields, it is also interesting to note that the ages compare well between the GMI-CTM and the GEOS-CTM, despite the fact that they are constrained with time-averaged and instantaneous fields, respectively. This suggests that the young mean ages associated with the RAna simulation more likely reflect the difference between using unbalanced analysis fields versus balanced assimilated fields and not to differences related to using averaged versus instantaneous fields. Indeed, further comparisons of the RAs3 simulation with a simulation constrained with three-hourly instantaneous assimilated fields reveal very small age differences (not shown).

Comparisons of vertical profiles of Γ_{GLB} (Figure 3b,c) show that the age differences are still larger in the middle stratosphere. Values of the stratospheric mean age in the RAna simulation differ from the CTM simulations by about ~ 30% over SH midlatitudes and by ~ 20% over NH midlatitudes. Note that, while the values of Γ_{GLB} for the RAs3 and RAs6 are somewhat less than those of the GMI CTM, they are much closer to the observations over the NH (where they are available), compared to the RAna simulation. Furthermore, the differences between the RAs3 and RAs6 simulations are very small, indicating that the mean age is less sensitive to differences in the IAU replay cycle frequency, compared to the difference between using assimilated versus analysis winds.

3.2 Comparison of an Oscillating Period χ_{tape}

The smaller values of Γ_{GLB} in the RAna simulation could be related to differences in tropical ascent among the simulations, with faster ascent out of the tropics resulting in overall younger stratospheric mean ages. We examine this further by comparing the vertical propagation of the tape recorder tracer χ_{tape} (Figure 4a). Comparison of the phase lag τ_ω and amplitude A of the propagating oscillation among the replay simulations reveals that the RAna simulation propagates annual oscillations too rapidly in the vertical and underattentuates the signal relative to both the RAs3 and RAs6 simulations and the observations (Figure 4b,c). This spurious vertical motion in the tropical lower and middle stratosphere is consistent with the findings in *Schoeberl et al.* [2003] and *Pawson et al.* [2007] using instantaneous six-hourly fields.

Alternatively, it is possible that the low stratospheric mean age values in the RAna simulation could be related to differences in mixing between the tropics and midlatitudes, with the in-mixing of (older) extratropical air into the tropics and (younger) tropical air into the extratropics. *Tan et al.* [2004] showed that the transport resulting from assimilated winds exhibits much larger mixing and entrainment rates, compared to those derived from general circulation model winds, and results from the generation of upward propagating features forced directly by the analysis increments during the assimilation process. Comparison of the meridional gradients in Γ_{GLB}, however, reveal more or less similar gradients among the simulations (not shown), indicating that the younger values of Γ_{GLB} in the RAna simulation are most likely not related to spurious mixing of midlatitude and tropical air masses.

We pursue the last point further by looking more systematically at transport signatures of mixing between low and high latitudes. Probability density functions of daily values of Γ_{GLB}, evaluated at 50 mb and averaged over 40°S-80°S during months in JJA, provide a sense for how well the polar vortex is isolated in the Southern Hemisphere [*Strahan and Polansky*, 2006]. While the degree of vortex isolation varies from month-to-month, all of the replay simulations capture the bimodal mean age pdfs characteristic of separated low latitude (young age) and high latitude (old age) air masses (Figure 5). Moreover, the widths of the pdfs differ little among the simulations, with only slightly wider pdfs for the simulations constrained with the assimilated fields. Thus, while the average values of the pdfs are different, their similar widths suggest that mixing over middle and high latitudes does not differ substantially among the replay simulations. This indicates that the young values of Γ_{GLB} in the RAna simulation are most likely related to spurious vertical transport out of the tropical lower stratosphere and not to differences in mixing between the tropics and extratropics.

3.3 Comparison of the Stratospheric Residual Circulation Ψ^*

The previous section showed that differences in vertical transport in the tropical lower stratosphere are most likely related to the large mean age differences among the replay simulations. In order to better understand the underlying differences in vertical transport we next compare the residual mean circulation $\overline{\Psi^*}$, which we define in terms of the Transformed Eulerian Mean [*Andrews et al.*, (1987]). The 2000-2010 climatological mean residual streamfunction $\overline{\Psi^*}$, shown for MERRA-2 in Figure 6 (panel a) features strong upwelling centered about the tropics, with a relatively stronger northern cell during December-January-February (DJF) and a stronger southern cell during June-July-August (JJA). A comparison of differences in upwelling, inferred from the vertical component of the residual mean velocity $\overline{w^*}$, averaged over 10°S-10°N, show consistently stronger tropical upwelling in the RAna simulation during both boreal winter and summer (panel b), relative both to MERRA-2 and to the RAs3 and RAs6 simulations. The differences in tropical upwelling are largest during JJA and on the order of 40% at 100 hPa, above which they decrease with increasing altitude.

The differences in $\overline{w^*}$ among the simulations may be understood in the context of previous work that has examined the relationship between the time-averaged stratospheric residual circulation and temperature increments in assimilated products. As discussed in *Weaver*

et al. [1993] there is no apriori reason to expect that a correct balance exists between the thermal and velocity fields at the time an observation is ingested during the assimilation cycle. This imbalance can excite unwanted intertial-gravity wave modes that affect the assimilated variables and manifest strongly in the residual vertical winds. As such, *Weaver et al.* [1993] was the first study to identify temperature increments as the cause of spurious vertical transport in the assimilation process.

Based on the findings in *Weaver et al.* [1993] we examine the extent to which differences in $\overline{w^*}$ reflect differences in the temperature tendency equations among the replay simulations (Figure 7). We first consider the 2000-2010 climatological mean balance of the temperature tendencies $\partial T/\partial t|_{PHYS}$ and $\partial T/\partial t|_{ANA}$ in the RAna simulation (panel a), where $\partial T/\partial t|_{ANA}$ denotes the analysis tendency introduced during the corrector segment of the IAU cycle (K/s) while $\partial T/\partial t|_{PHYS}$ is the total diabatic temperature tendency (K/s) resulting from moist, radiative, gravity wave and turbulent processes. In the annual climatological mean, the sum of $\partial T/\partial t|_{PHYS}$ and $\partial T/\partial t|_{ANA}$ roughly balance $\partial T/\partial t|_{DYN}$ so that the relative contributions of the physics and analysis tendencies to the vertical mass flux can be evaluated.

Consistent with larger values of upwelling $\overline{w^*}$ in the tropics (Figure 6b) the sum $\partial T/\partial t|_{PHYS}$ + $\partial T/\partial t|_{ANA}$ is larger in the SH subtropics in the RAna simulation, compared to both the RAs3 and RAs6 simulations (Fig. 7b). Further comparison of the individual components $\partial T/\partial t|_{PHYS}$ and $\partial T/\partial t|_{ANA}$ among the simulations reveals that the large differences in the implied mass flux among the simulations are related to differences in the analysis tendency $\partial T/\partial t|_{ANA}$, which is largest in the RAna simulation in the tropics where the underlying GCM features a cold temperature bias (Fig. 7d). By comparison, the differences in $\partial T/\partial t|_{PHYS}$ among the simulations is much smaller (Fig. 7c). This suggests that the $\overline{w}^*$ differences are compensating for differences in the temperature increments, which are largest when the replay technique is applied to the instantaneous analysis fields. In other words, simulations constrained with the analysis fields produce larger temperature increments that require larger (compensating) upwelling in the tropics.

4 Large-Scale Tropospheric Transport

While the use of the time-averaged assimilated meteorological fields helps suppress spurious vertical transport in the lower stratosphere, it is not clear what (if any) impact this has on large-scale climatological mean transport in the troposphere. Furthermore, it is possible that the use of time-averaged assimilated fields may smear out discontinuities at fronts, although not to the same extent as in a CTM, wherein the meteorology is only updated at the analysis frequency (i.e. not applied as a forcing at every model time step). To address these issues we now examine differences in the climatological distributions of the tropospheric tracers among the RAna, RAs3 and RAs6 simulations, beginning with the 5-day and 50-day loss tracers, $\overline{\chi}_5$ and $\overline{\chi}_{50}$ (Figure 8a,b).

In all of the simulations the patterns of $\overline{\chi}_5$ and $\overline{\chi}_{50}$ decrease poleward away from the midlatitude source region Ω_{MID}, consistent with the findings in *Orbe et al.* [2016] for a coarser resolution version of the GMI-CTM driven with MERRA meteorological fields. The transport differences among the simulations are negligible, with only a weak indication of more rapid transport northward in the RAna simulation, manifest in the 50-day lifetime tracer over northern high latitudes $\overline{\chi}_{50}$ (Fig. 8b). This difference in $\overline{\chi}_{50}$ among the simulations also occurs in a region where there are slightly larger values of $\overline{\chi}_{\mathrm{STE}}$ in the RAna simulation (Fig. 8c) suggesting slightly more rapid isentropic mixing in the RAna simulation in the NH upper troposphere. Note, however, that the values of χ_{STE} in the lower troposphere are extremely small compared to near the tropopause, where the simulations are nearly identical, indicating only negligible differences in stratosphere-troposphere exchange. The differences in the mean age Γ_{NH} over SH high latitudes, while not negligible, are also small (Fig. 8d).

To put the above differences in context we compare profiles of two of the tropospheric tracers (χ_5 and Γ_{NH}) among the replay simulations with profiles from a variety of other simulations (Figure 9). These include the original GMI-CTM and GEOS-CTM simulations (first two rows, Table 1), as well as another integration of the GMI-CTM driven with MERRA-2 meteorological fields run at a coarser (2° latitude x 2.5° longitude) resolution. Two additional GMI-CTM simulations constrained with MERRA [*Rienecker et al.*, 2011] at both 2° latitude x 2.5° longitude and 1° latitude x 1.5° longitude resolutions are also shown to illustrate the impacts of transport differences related to differences in MERRA versus MERRA-2. Finally, we also include the results from two simulations of the Whole Atmosphere Community Climate Model (WACCM) [*Marsh et al.*, 2013] nudged to MERRA temperature, surface pressure, and zonal meridional winds on relaxation timescales of 5 hours and 50 hours using the approach of *Kunz et al.* [2011]. The GMI-CTM 2° simulation constrained by MERRA fields as well as the WACCM MERRA simulations were presented in *Orbe et al.* [2017].

Comparisons of free tropospheric profiles of $\overline{\chi_5}$ and the NH midlatitude mean age $\overline{\Gamma}_{\mathrm{NH}}$ in the Northern Hemisphere and Southern Hemisphere, respectively, show that the largest differences generally occur between the NASA models (i.e. GMI and GEOS-5 Replay) and the WACCM model simulations. Note that even though there are some differences between both the CTM and GEOS-5 replay simulations, manifest in upper troposphere values of $\overline{\chi}_5$ at 20°N and 50°N (Fig. 9, bottom panels), these differences are mainly confined to the upper troposphere and lower stratosphere. Furthermore, they are much smaller than the differences in both $\overline{\chi_5}$ and $\overline{\Gamma}_{\mathrm{NH}}$ between the NASA versus WACCM simulations. As discussed in *Orbe et al.* [2017] these differences in large-scale transport among the NASA and WACCM simulations are most likely related to differences in parameterized convection. Thus, the transport differences related to using time-averaged versus instantaneous winds are relatively negligible in the troposphere, compared to the differences in (parameterized) convective transport. This is consistent with previous studies linking large differences in interhemispheric transport [*Denning et al.*, 1999] and methane lifetimes [*Patra et al.*, 2011] to differences in vertical mixing by parameterized convection.

5 Conclusions

The GEOS replay technique represents a powerful tool for performing atmospheric composition studies that complements offline chemical transport models. Unlike CTMs, however, all of the subgrid-scale processes are recalculated by the model so that they are consistent with the (resolved) analysis fields, which can be important for chemical species that are sensitive to the parameterized components of the flow and/or when performing coarse resolution simulations.

The fidelity of GEOS replay simulations for representing atmospheric composition rests on their ability to produce credible atmospheric transport. The main goal of this study, therefore, has been to document the large-scale transport properties of various GEOS-5 replay simulations using idealized tracers that probe the transport circulation on timescales ranging from a few days to several years. Comparisons among simulations constrained with instantaneous analysis fields versus time-averaged assimilated fields reveal large differences in stratospheric transport. Our findings are as follows:

- Replay simulations using six-hourly instantaneous analysis fields are characterized by stratospheric mean ages (Γ_{GLB}) that are at places 30% too young, relative to both two offline models and to observations. By comparison, replay simulations constrained with the time-averaged assimilated fields produce more credible stratospheric transport. Furthermore, the largest stratospheric transport difference are associated with whether assimilated versus analysis fields are used, and not to the temporal filtering of the constraining fields (i.e. time-averaged versus instantaneous).
- Smaller values of Γ_{GLB} in the RAna simulation are consistent with stronger tropical upwelling in the lower stratosphere that acts in response to large temperature analysis

increments. This indicates that the mean age associated with the analysis is too young, relative to both the observations and the smoother more dynamically balanced time-averaged assimilated fields.

- The climatological mean large-scale tropospheric transport properties are relatively insensitive to how the flow is specified when the replay technique is used. Passive tracer-based measures of interhemispheric transport, transport to the Arctic and stratosphere-troposphere-exchange differ negligibly depending on whether instantaneous or time-averaged winds are used, relative to the differences in convective transport associated with different convective parameterizations.

One caveat in this study is that our focus has been on large-scale climatological mean transport since our primary interest has been in validating the replay technique for its use in long chemistry-climate simulations, such as those conducted for CCMI. We do not discuss transport variability on daily and/or regional spatial scales, in which case the subtleties of using time-averaged versus instantaneous winds in the troposphere may be important. While preliminary analyses of regional transport differences indicate that they are no larger than 5%, this is not the focus of this study and we reserve further examination for future work.

We close by briefly mentioning other applications of the replay technique with the GEOS model. Coupled atmosphere-ocean experiments, in which the large-scale atmospheric meteorological fields are constrained using replay, are currently used to generate the ocean data assimilation. Replay simulations can also be used to study certain aspects of feedbacks that are not currently studied in offline models. For example, replay simulations can be used to study the impacts of stratospheric chemistry on water vapor, which is prescribed from a climatology in the CTM, but can be treated as an online variable in a replay simulation. [Note that this represents only a limited assessment of feedbacks since the meteorological fields are still constrained to the analysis (and thus not responding to the chemistry as in a fully online simulation).] Among others, these applications of the replay technique will need to be separately validated in future studies.

Acknowledgments

The authors thank the high-performance computing resources provided by NASA's Advanced Supercomputing (NAS) Division and the NASA Center for Climate Simulation (NCCS) as well as NASA's Modeling, Analysis and Prediction (MAP) program, which supports the Global Modeling Assimilation Office and core chemistry-climate and chemistry-modeling activities. They also thank Jean-Francois Lamarque, Simone Tilmes and Douglas E. Kinnison for providing the WACCM data. WACCM is a component of the Community Earth System Model (CESM), which is supported by the National Science Foundation (NSF) and the Office of Science of the U.S. Department of Energy. Computing resources were provided by NCAR's Climate Simulation Laboratory, sponsored by NSF and other agencies. D.W. acknowledges support from NSF grant AGS-1403676 and NASA grant NNX14AP58G. The GEOS-5 source code used in this study is available under the NASA Open Source Agreement at http://opensource.gsfc.nasa.gov/projects/GEOS-5/. All model output is available for public access under ftp://gmaoftp.gsfc.nasa.gov/pub/data/corbe/Replay. The SF_6- and CO_2- in situ stratospheric mean age data can be accessed by contacting Andreas Engel (an.engel@iau.uni-frankfurt.de).

References

Andrews, A., K. Boering, B. Daube, S. Wofsy, M. Loewenstein, H. Jost, J. Podolske, C. Webster, R. Herman, D. Scott, et al. (2001), Mean ages of stratospheric air derived from in situ observations of co2, ch4, and n2o, *Journal of Geophysical Research: Atmospheres*, *106*(D23), 32,295–32,314.

Andrews, D., J. Holton, and C. Leovy ((1987)), Middle Atmosphere Dynamics, *Academic Press*, *60*, 489, doi:10.1175/1520-0469(2003)060<0103:CEOOAL>2.0.CO;2.

Bloom, S., L. Takacs, A. Da Silva, and D. Ledvina (1996), Data assimilation using incremental analysis updates, *Monthly Weather Review*, *124*(6), 1256–1271.

Boering, K., S. Wofsy, B. Daube, H. Schneider, M. Loewenstein, J. Podolske, and T. Conway (1996), Stratospheric transport rates and mean age distribution derived from observations of atmospheric co2 and n2o, *Science*, *274*(1), 340–1343.

Colarco, P., A. da Silva, M. Chin, and T. Diehl (2010), Online simulations of global aerosol distributions in the NASA GEOS-4 model and comparisons to satellite and ground-based aerosol optical depth, *Journal of Geophysical Research: Atmospheres*, *115*(D14).

Denning, A. S., M. Holzer, K. R. Gurney, M. Heimann, R. M. Law, P. J. Rayner, I. Y. Fung, S.-M. Fan, S. Taguchi, P. Friedlingstein, et al. (1999), Three-dimensional transport and concentration of sf6 a model intercomparison study (transcom 2), *Tellus B: Chemical and Physical Meteorology*, *51*(2), 266–297.

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

Eyring, V., J.-F. Lamarque, P. Hess, F. Arfeuille, K. Bowman, M. P. Chipperfiel, B. Duncan, A. Fiore, A. Gettelman, M. A. Giorgetta, et al. (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.

Geller, L., J. Elkins, L. JM, A. Clarke, D. Hurst, J. Butler, and R. Myers (1997), Tropospheric SF6: Observed latitudinal distribution and trends, derived emissions and interhemispheric exchange time, *Geophys Res Lett*, *24*, 675–678.

Hall, T. M., D. W. Waugh, K. A. Boering, and R. A. Plumb (1999), Evaluation of transport in stratospheric models, *Journal of Geophysical Research: Atmospheres*, *104*(D15), 18,815–18,839.

Kouatchou, J., A. Molod, J. Nielsen, B. Auer, P. W, and T. Clune (2015), Geos-5 chemistry transport model user's guide.

Kunz, A., L. Pan, P. Konopka, D. Kinnison, and S. Tilmes (2011), Chemical and dynamical discontinuity at the extratropical tropopause based on START08 and WACCM analyses, *Journal of Geophysical Research: Atmospheres*, *116*(D24).

Lamarque, J., L. Emmons, P. Hess, D. E. Kinnison, S. Tilmes, F. Vitt, C. Heald, E. A. Holland, P. Lauritzen, J. Neu, et al. (2012), Cam-chem: Description and evaluation of interactive atmospheric chemistry in the Community Earth System Model, *Geosci. Model Dev*, *5*(2), 369–411.

Legras, B., I. Pisso, G. Berthet, and F. Lefevre (2004), Variability of the lagrangian turbulent diffusivity in the lower stratosphere, *Atmospheric Chemistry and Physics Discussions*, *4*(6), 8285–8325.

Levin, I., and V. Hesshaimer (1996), Refining of atmospheric transport model entries by the globally observed passive tracer distributions of 85krypton and sulfur hexafluoride (SF6), *Journal of Geophysical Research: Atmospheres (1984–2012)*, *101*(D11), 16,745–16,755.

Marsh, D. R., M. J. Mills, D. E. Kinnison, J.-F. Lamarque, N. Calvo, and L. M. Polvani (2013), Climate change from 1850 to 2005 simulated in CESM1 (WACCM), *Journal of climate*, *26*(19), 7372–7391.

Meijer, E. W., B. Bregman, A. Segers, and P. F. van Velthoven (2004), The influence of data assimilation on the age of air calculated with a global chemistry-transport model using ecmwf wind fields, *Geophysical research letters*, *31*(23).

Molod, A., L. Takacs, M. Suarez, and J. Bacmeister (2015), Development of the geos-5 atmospheric general circulation model: evolution from merra to merra2, *Geoscientific Model Development*, *8*(5), 1339–1356.

Monge-Sanz, B., M. Chipperfield, D. Dee, A. Simmons, and S. Uppala (2013), Improvements in the stratospheric transport achieved by a chemistry transport model with ecmwf (re) analyses: Identifying effects and remaining challenges, *Quarterly Journal of the Royal*

Meteorological Society, *139*(672), 654–673.

Moorthi, S., and M. J. Suarez (1992), Relaxed Arakawa-Schubert. A parameterization of moist convection for general circulation models, *Monthly Weather Review*, *120*(6), 978–1002.

Mote, P. W., K. H. Rosenlof, M. E. McIntyre, E. S. Carr, J. C. Gille, J. R. Holton, J. S. Kinnersley, H. C. Pumphrey, J. M. Russell III, and J. W. Waters (1996), An atmospheric tape recorder: The imprint of tropical tropopause temperatures on stratospheric water vapor.

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

Orbe, C., D. W. Waugh, P. A. Newman, and S. Steenrod (2016), The transit-time distribution from the Northern Hemisphere midlatitude surface, *Journal of the Atmospheric Sciences*, (2016).

Orbe, C., D. W. Waugh, H. Yang, L. J. F., S. Tilmes, and D. E. Kinnison (2017), Tropospheric transport differences between models using the same large-scale meteorological fields, *Geophysical Research Letters*.

Patra, P. K., S. Houweling, M. Krol, P. Bousquet, D. Belikov, D. Bergmann, H. Bian, P. Cameron-Smith, M. P. Chipperfield, K. Corbin, et al. (2011), Transcom model simulations of ch 4 and related species: linking transport, surface flux and chemical loss with ch 4 variability in the troposphere and lower stratosphere, *Atmospheric Chemistry and Physics*, *11*(24), 12,813–12,837.

Pawson, S., I. Stajner, S. R. Kawa, H. Hayashi, W.-W. Tan, J. E. Nielsen, Z. Zhu, L.-P. Chang, and N. J. Livesey (2007), Stratospheric transport using 6-h-averaged winds from a data assimilation system, *Journal of Geophysical Research: Atmospheres*, *112*(D23).

Prather, M. J., X. Zhu, S. E. Strahan, S. D. Steenrod, and J. M. Rodriguez (2008), Quantifying errors in trace species transport modeling, *Proceedings of the National Academy of Sciences*, *105*(50), 19,617–19,621.

Prather, M. J., X. Zhu, Q. Tang, J. Hsu, and J. L. Neu (2011), An atmospheric chemist in search of the tropopause, *Journal of Geophysical Research: Atmospheres*, *116*(D4).

Putnam, W. M., and S.-J. Lin (2007), Finite-volume transport on various cubed-sphere grids, *Journal of Computational Physics*, *227*(1), 55–78.

Rasch, P., N. Mahowald, and B. Eaton (1997), Representations of transport, convection, and the hydrologic cycle in chemical transport models: Implications for the modeling of short-lived and soluble species, *Journal of Geophysical Research: Atmospheres*, *102*(D23), 28,127–28,138.

Rienecker, M., M. J. Suarez, R. Todling, J. Bacmeister, L. Takacs, H. Liu, W. Gu, M. Sienkiewicz, R. Koster, R. Gelaro, I. Stajner, and J. E. Nieselen (2008), The GEOS-5 Data Assimilation System-Documentation of Versions 5.0. 1, 5.1. 0, and 5.2. 0. technical report series on global modeling and data assimilation, *27*, 1–118.

Rienecker, M. M., M. J. Suarez, R. Gelaro, R. Todling, J. Bacmeister, E. Liu, M. G. Bosilovich, S. D. Schubert, L. Takacs, G.-K. Kim, et al. (2011), MERRA: NASA's modern-era retrospective analysis for research and applications, *Journal of Climate*, *24*(14), 3624–3648.

Schoeberl, M. R., A. R. Douglass, Z. Zhu, and S. Pawson (2003), A comparison of the lower stratospheric age spectra derived from a general circulation model and two data assimilation systems, *Journal of Geophysical Research: Atmospheres*, *108*(D3).

Strahan, S., and B. Polansky (2006), Meteorological implementation issues in chemistry and transport models, *Atmospheric Chemistry and Physics*, *6*(10), 2895–2910.

Strahan, S., B. Duncan, and P. Hoor (2007), Observationally derived transport diagnostics for the lowermost stratosphere and their application to the GMI chemistry and transport model, *Atmospheric Chemistry and Physics*, *7*(9), 2435–2445.

Strahan, S., A. Douglass, and S. Steenrod (2016), Chemical and dynamical impacts of stratospheric sudden warmings on arctic ozone variability, *Journal of Geophysical Research: Atmospheres*, *121*(19).

Strode, S., B. Duncan, E. Yegorova, J. Kouatchou, J. Ziemke, and A. Douglass (2015), Implications of carbon monoxide bias for methane lifetime and atmospheric composition in chemistry climate models, *Atmospheric Chemistry and Physics*, *15*(20), 11,789–11,805.

Strode, S. A., H. M. Worden, M. Damon, A. R. Douglass, B. N. Duncan, L. K. Emmons, J.-F. Lamarque, M. Manyin, L. D. Oman, J. M. Rodriguez, et al. (2016), Interpreting space-based trends in carbon monoxide with multiple models, *Atmospheric Chemistry and Physics*, *16*(11), 7285–7294.

Tan, W. W., M. A. Geller, S. Pawson, and A. Da Silva (2004), A case study of excessive subtropical transport in the stratosphere of a data assimilation system, *Journal of Geophysical Research: Atmospheres*, *109*(D11).

Thiele, G. and J.L. Sarmiento (1990),Tracer dating and ocean ventilation, *Journal of Geophysical Research: Oceans*, *95*(C6), 9377–9391.

Volk, C., J. Elkins, D. Fahey, R. Salawitch, et al. (1996), Quantifying transport between the tropical and mid-latitude lower stratosphere, *Science*, *272*(5269), 1763.

Waugh, D., T. Hall, W. Randel, P. Rasch, B. Boville, K. Boering, S. Wofsy, B. Daube, J. Elkins, D. Fahey, et al. (1997), Three-dimensional simulations of long-lived tracers using winds from maccm2, *Journal of Geophysical Research: Atmospheres*, *102*(D17), 21,493–21,513.

Waugh, D., A. Crotwell, E. Dlugokencky, G. Dutton, J. Elkins, B. Hall, E. Hintsa, D. Hurst, S. Montzka, D. Mondeel, et al. (2013), Tropospheric SF6: Age of air from the Northern Hemisphere midlatitude surface, *Journal of Geophysical Research: Atmospheres*, *118*(19), 11–429.

Weaver, C. J., A. R. Douglass, and R. B. Rood (1993), Thermodynamic balance of three-dimensional stratospheric winds derived from a data assimilation procedure, *Journal of the atmospheric sciences*, *50*(17), 2987–2993.

Yu, K., C. A. Keller, D. J. Jacob, A. M. Molod, S. D. Eastham, and M. S. Long (2017), Errors and improvements in the use of archived meteorological data for chemical transport modeling, *Geosci. Model Dev. Discuss.*, *2017*, 1–22.

Ziemke, J., M. Olsen, J. Witte, A. Douglass, S. Strahan, K. Wargan, X. Liu, M. Schoeberl, K. Yang, T. Kaplan, et al. (2014), Assessment and applications of nasa ozone data products derived from aura omi/mls satellite measurements in context of the gmi chemical transport model, *Journal of Geophysical Research: Atmospheres*, *119*(9), 5671–5699.

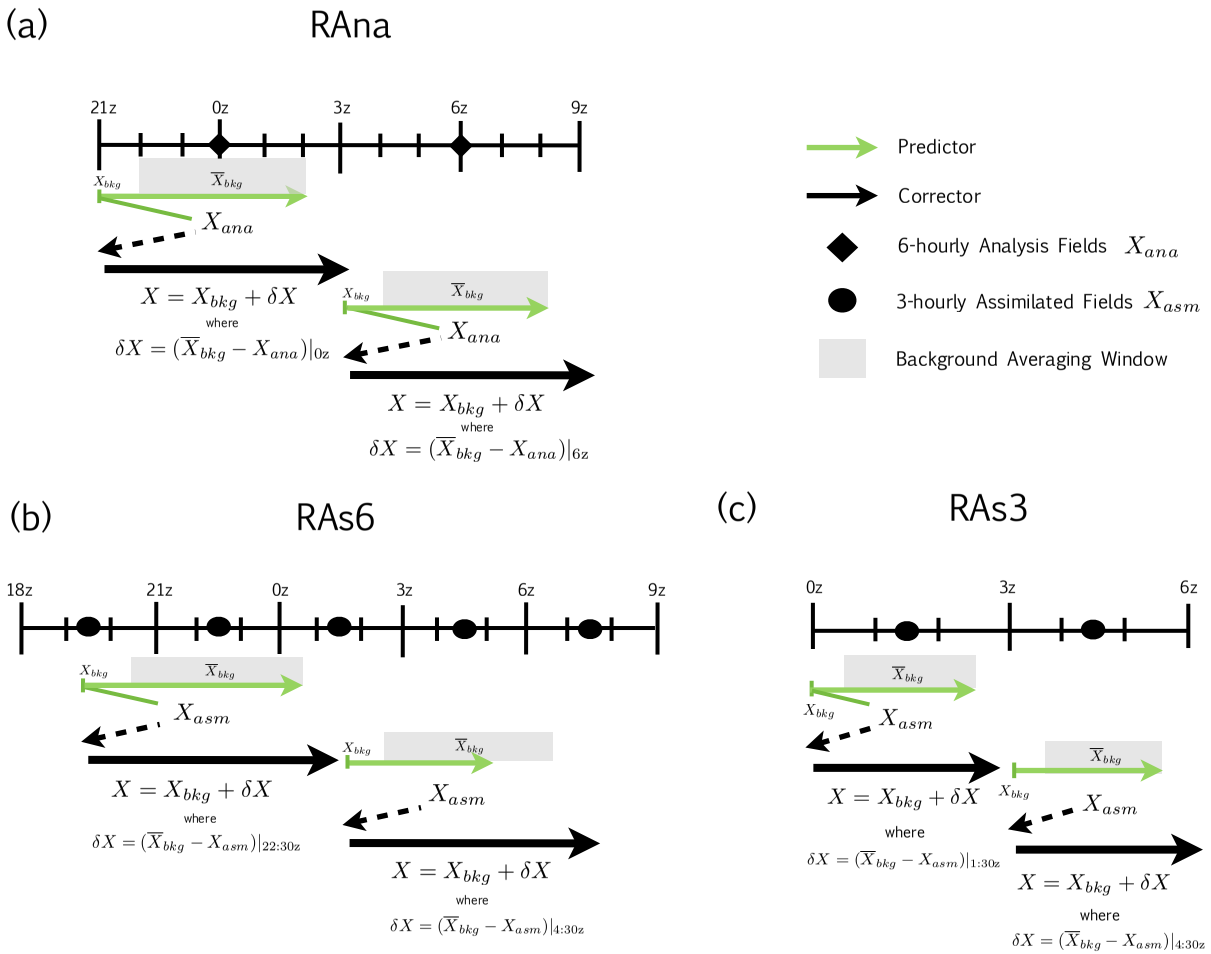

Figure 1. Schematic illustrations of the replay simulations examined in this study. The replay technique uses the same Incremental Analysis Update (IAU) technique [*Bloom et al.*, 1996] that was used to generate the MERRA-2 data assimilation, which consists of both predictor segments (green lines) and corrector segments (black lines). For the RAna simulation (panel a) a five-hour forecast centered about the analysis time (0z) is launched at 21z; an increment δX is then calculated as the difference between the time-averaged background state $\overline{X}_{bkg}|_{0z}$ centered about 0z and a pre-existing analysis field X_{ana} (black diamonds). The model is then backtracked to 21z and the increment δX is linearly applied to the background state $X_{bkg}|_{21z}$ over a six hour corrector interval. For the RAs6 (panel b) and RAs3 (panel c) simulations the increment δX is calculated with respect to the three-hourly time-averaged MERRA-2 assimilated fields, which are centered about 01:30 GMT, 04:30 GMT, 07:30 GM, and so on (black circles). The replay corrector segment occurs ever six hours in RAs6 and every 3 hours in RAs3. The averaged background state used to calculate the increment δX in (a) and (b) is a linear four-hour time average (± 2 hours) that is used to partly suppress resonance frequencies shorter than the IAU six-hour cycle. A two-hour time average (± 1 hour) is used in (c), when the corrector segment spans three hours.

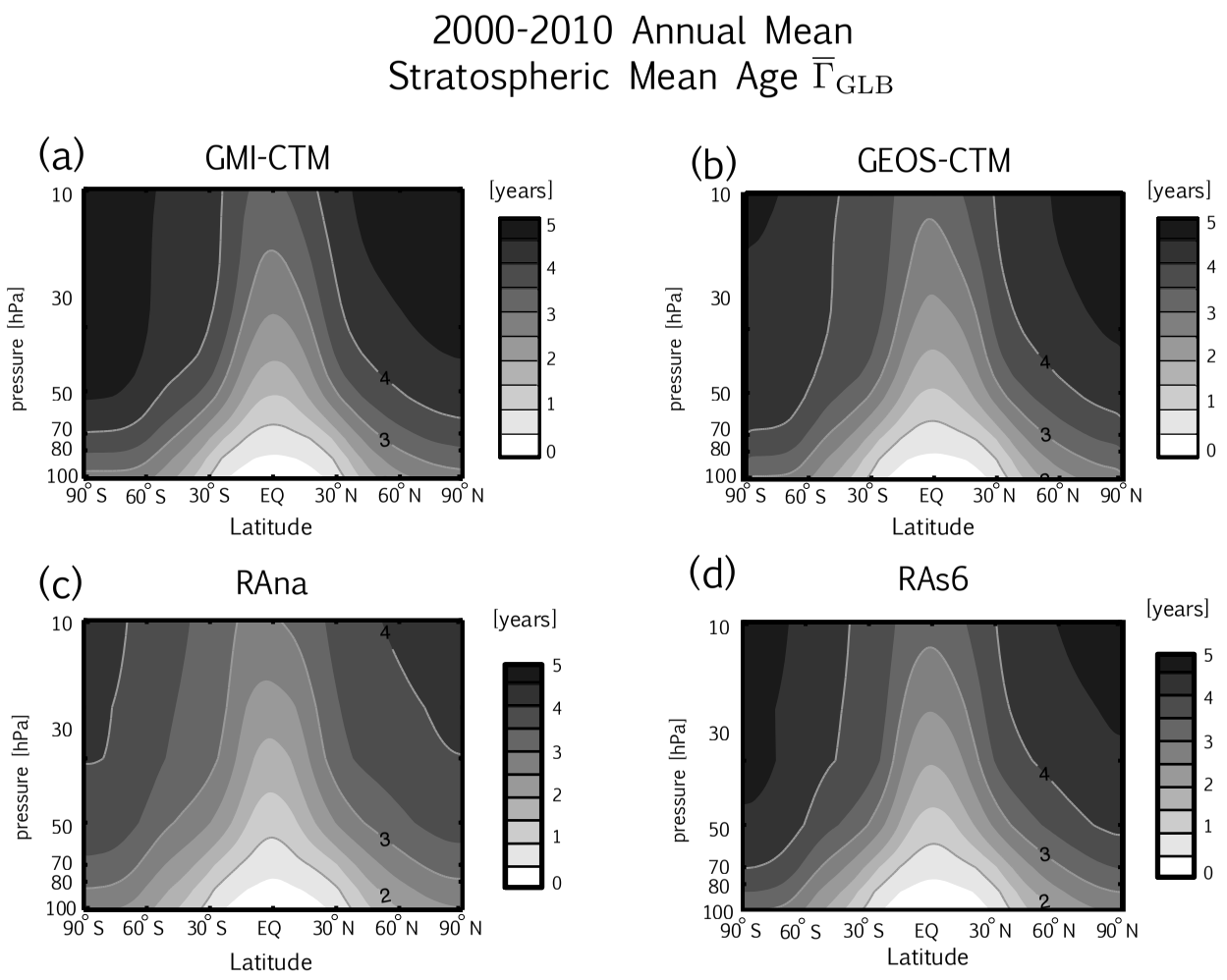

Figure 2. 2000-2010 climatological mean annually averaged mean age of air, Γ_{GLB}, in the stratosphere [contour spacing: 0.5 years]. Model output is shown for the GMI-CTM (a) and the GEOS-CTM (b), driven with time-averaged and instantaneous MERRA-2 3 hourly assimilated fields, respectively. c)-d) Same as in (a), except shown for the RAna and RA6 simulations, in which GEOS-5 is constrained by instantaneous six-hourly analysis fields and time-averaged three-hourly assimilated fields, respectively.

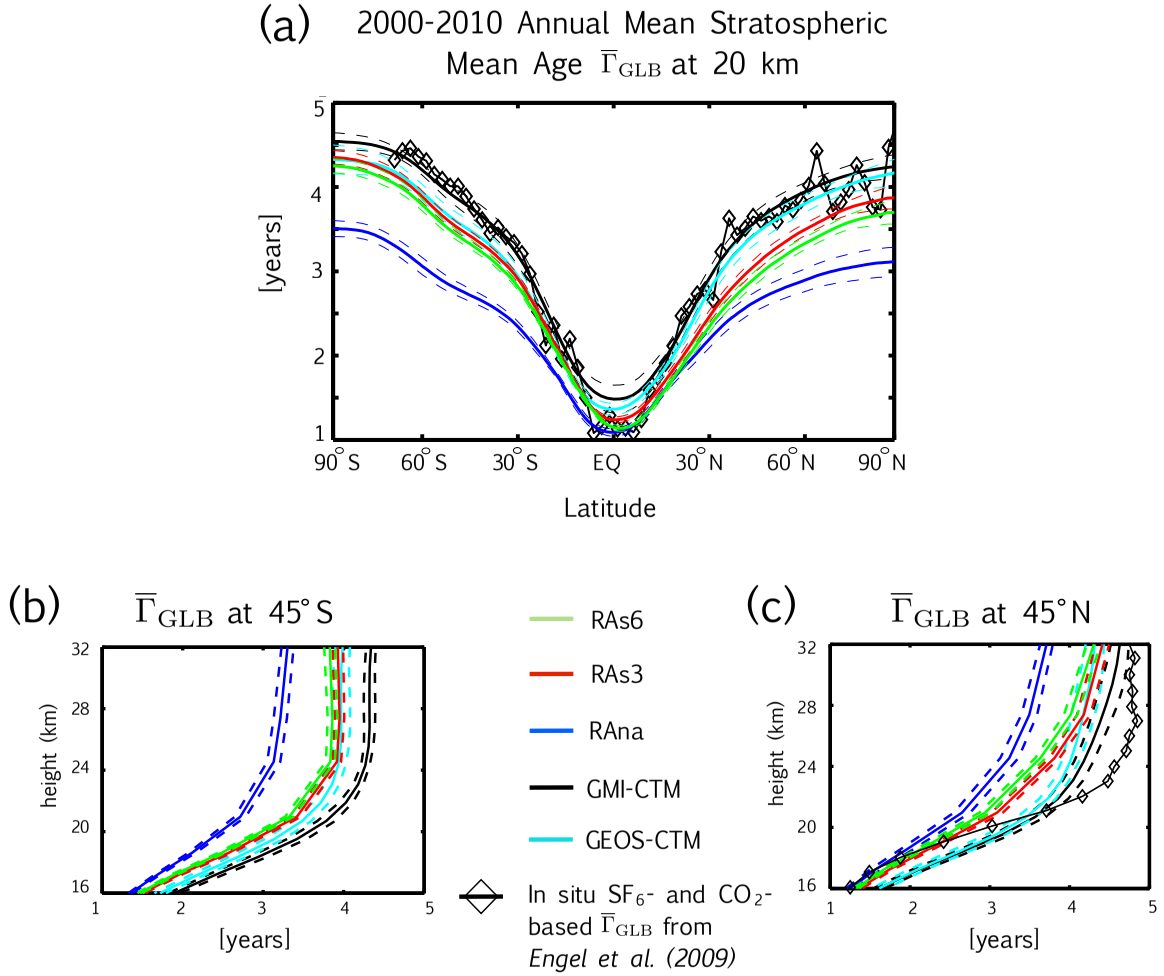

Figure 3. (a) Meridional profiles of the 2000-2010 climatological mean annually averaged mean age of air, Γ_{GLB}, evaluated at 20 km. (b-c) Vertical profiles of the climatological mean annually averaged mean age of air, Γ_{GLB}, evaluated at b) 45°S and c) 45°N. Output from the RAna, RAs6 and RAs3 simulations are shown in the blue, green and red lines, respectively. The black and cyan lines correspond to the GMI-CTM and GEOS-CTM simulations, respectively. Dashed lines denote $\pm\sigma_\Gamma$, the standard deviation of Γ_{GLB} over the 2000-2010 climatological averaging period. The diamonds correspond to SF_6 and CO_2 in situ based measures of the stratospheric mean age taken from *Boering et al.* [1996] and *Engel et al.* [2009]. Note that observational constraints on the vertical profiles of Γ_{GLB} are only available for the NH.

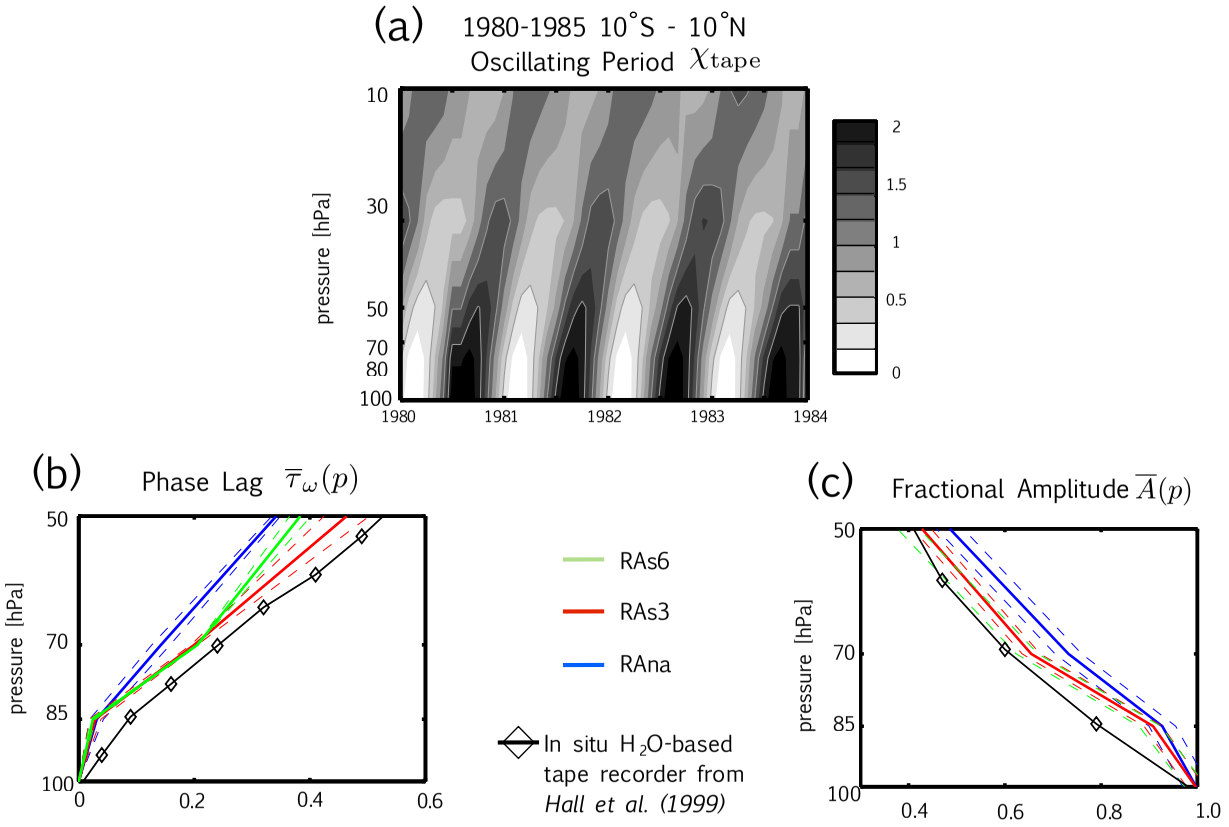

Figure 4. a) Evolution of the tropical oscillating period evaluated over 10°S and 10°N over the first four years of the RAna simulation. Note that the maximum at 100 mb over 10°S and 10°N has been specified to occur during October, when water vapor concentrations there are the largest [*Mote et al.*, 1996]. Equatorial profiles of the 1980-1985 climatological mean phase lag τ_ω (b) and the fractional amplitude A (c) of a propagating period relative to the tracer source region at 100 mb are also shown, averaged between 10°N and 10°S. The blue, green and red lines correspond to the RAna, RAs6 and RAs3 simulations, respectively. The diamonds denote observed values of τ_ω and the fractional amplitude A based on in situ water vapor measurements from HALOE [*Mote et al.*, 1996; *Hall et al.*, 1999].

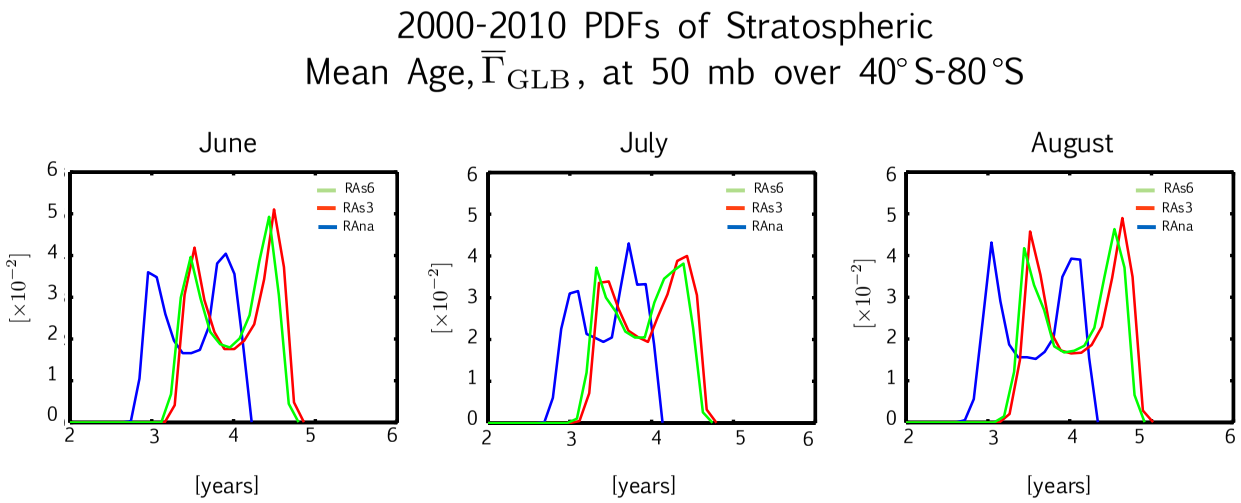

Figure 5. Monthly mean 2000-2010 climatological mean PDFs of daily stratospheric mean age Γ_{GLB} values evaluated over 40°S and 80°S at 50 mb for June (left), July (middle) and August (right). The RAna, RAs3 and RAs6 simulations are shown in the blue, green and red lines, respectively.

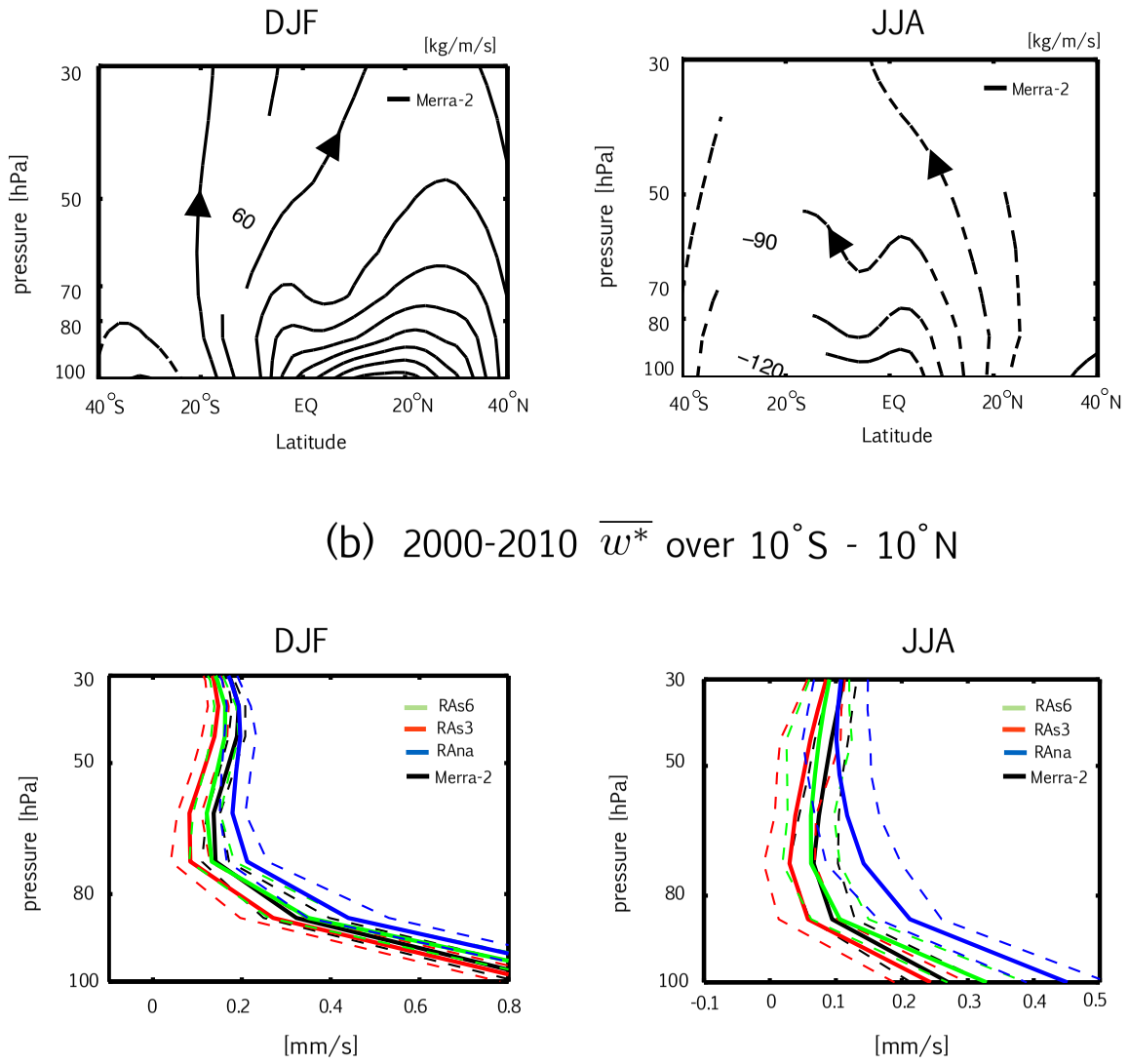

Figure 6. a) The boreal winter (DJF) (left) and boreal summer (JJA) 2000-2010 climatological mean residual streamfunction Ψ^*, approximated using the Transformed Eulerian mean (TEM). Output is shown for MERRA-2. Streamfunction contours are spaced 30 kg m^{-1} s^{-2} apart, with positive(negative) values depicted in solid(dashed) lines. (b) Comparison of the 2000-2010 DJF (left) and JJA (right) residual mean vertical velocity w^* averaged over latitudes spanning 10°S and 10°N. Dashed lines in (b) correspond to the standard deviations $\pm\sigma_{w^*}$ over the 2000-2010 climatological averaging period. The RAna, RAs3 and RAs6 simulations are shown in the blue, green and red lines, respectively.

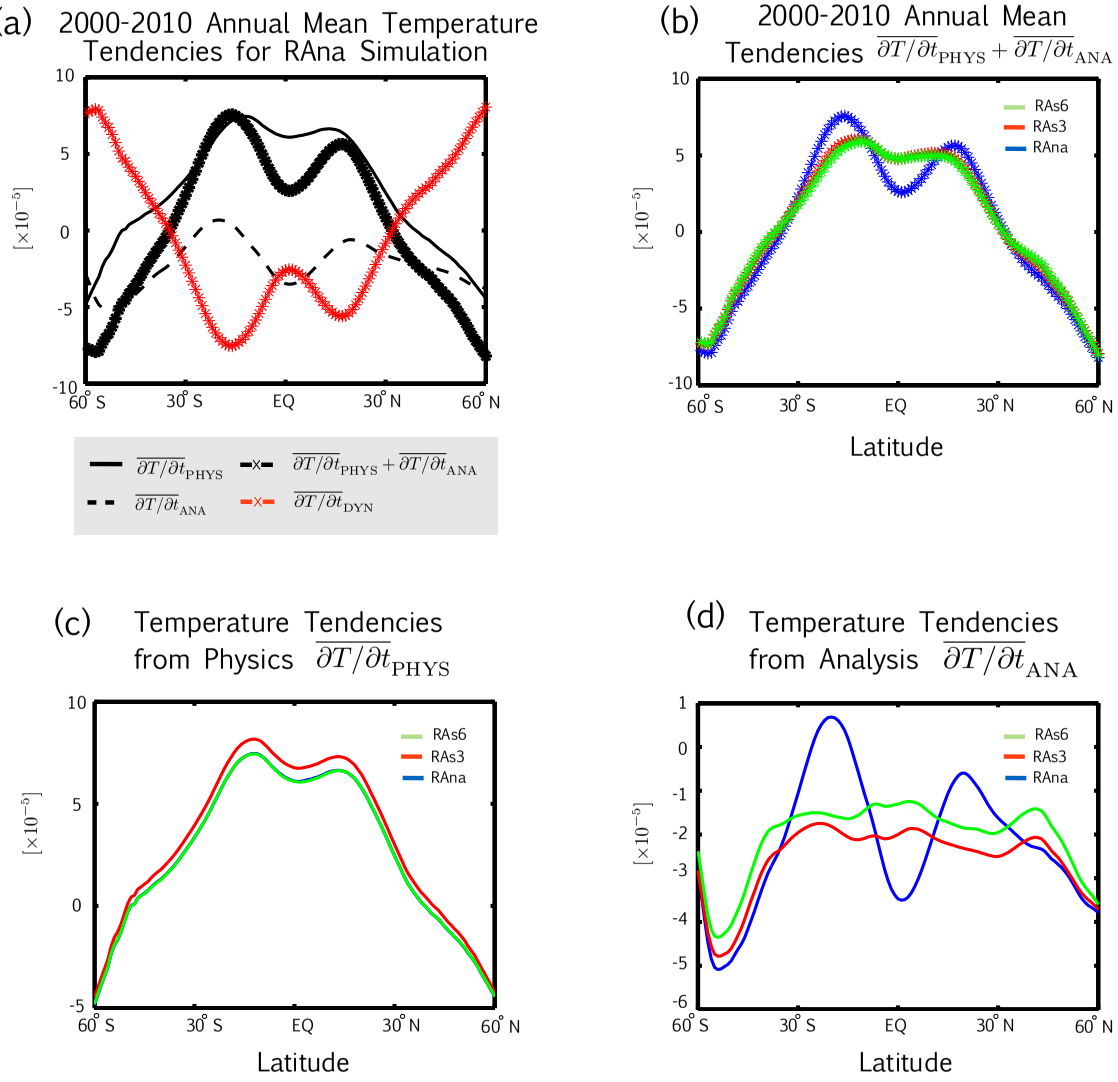

Figure 7. a) 2000-2010 annually averaged climatological mean temperature tendency budget for the RAna simulation evaluated at 85 mb. The different lines correspond to $\partial T/\partial t_{\rm PHYS}$, the total diabatic temperature tendency [K/s] due to physics (e.g. turbulence, moist processes, gravity wave drag, frictional dissipation, and radiation) (thin solid black line), $\partial T/\partial t_{\rm ANA}$, the analysis tendency introduced during the corrector segment of the IAU cycle (K/s) (thin dashed black line), the sum of $\partial T/\partial t_{\rm PHYS}$ and $\partial T/\partial t_{\rm ANA}$ (thick black starred line) and the temperature tendency due to dynamics, $\partial T/\partial t_{\rm DYN}$ (thick red starred line). Note that in the long term climatological mean the analysis and total physics tendencies roughly balance the dynamical tendency. (b) Comparison of the sum of $\partial T/\partial t_{\rm PHYS}$ and $\partial T/\partial t_{\rm ANA}$ among the RAna (blue), RAs6 (green) and RAs3 (red) simulations. Comparisons of the individual terms $\partial T/\partial t_{\rm PHYS}$ and $\partial T/\partial t_{\rm ANA}$ among the simulations are shown in (c) and (d), respectively.

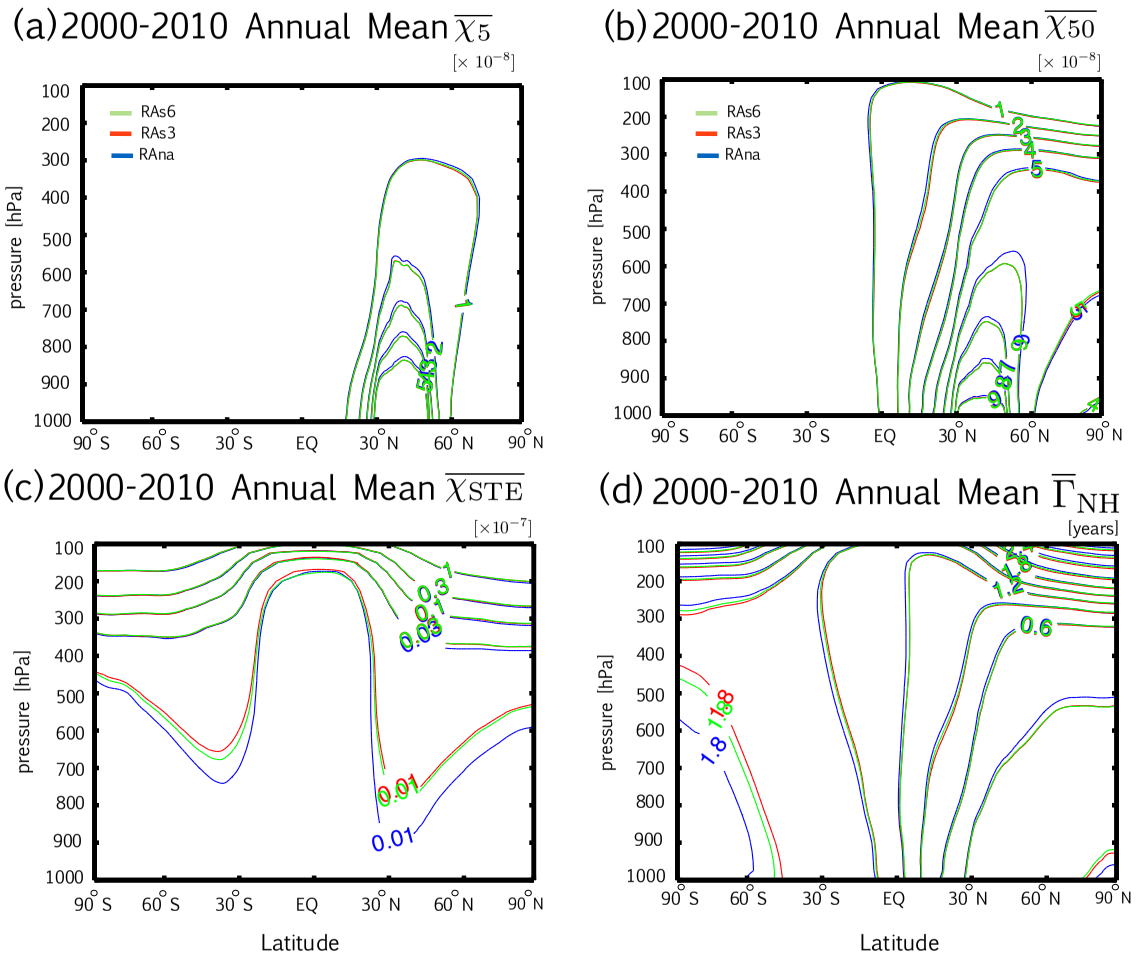

Figure 8. 2000-2010 annually averaged climatological mean distributions of the 5-day and 50-day lifetime tracers $\overline{\chi}_5$ (a) and $\overline{\chi}_{50}$ (b), the STE tracer, $\overline{\chi}_{\mathrm{STE}}$ (c) and the mean age since air last contacted the NH midlatitude surface $\overline{\Gamma}_{\mathrm{NH}}$ (d). As in previous figures, the blue, green and red lines denote the RAna, RAs6 and RAs3 simulations.

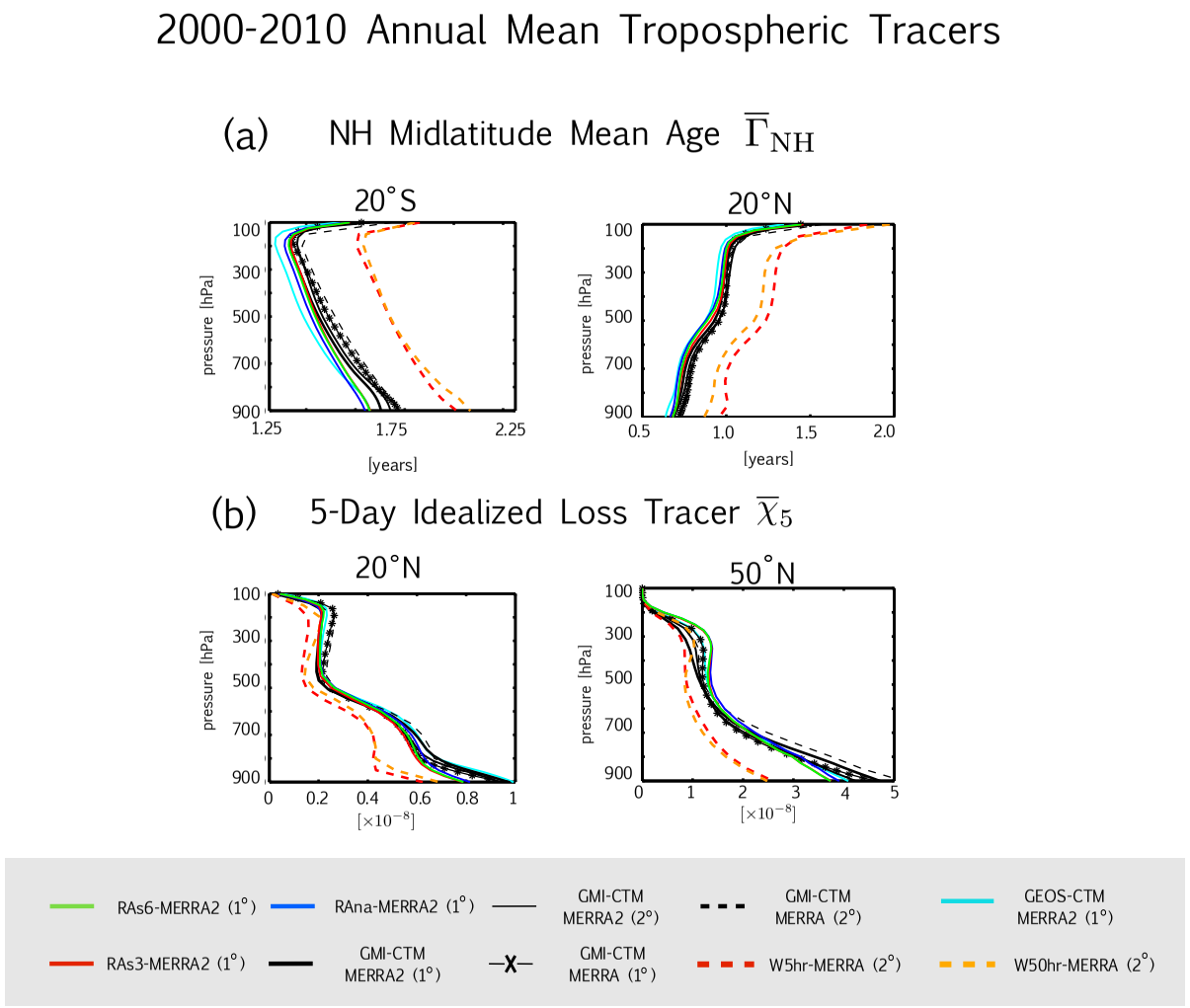

Figure 9. Vertical profiles of the annual mean climatological mean tracers $\overline{\Gamma}_{NH}$ (a), evaluated at 20°S (left) and 20°N (right), and $\overline{\chi}_5$ (b), evaluated at 20°N (left) and 50°N (right). As in previous figures, the blue, green, red (thin solid) lines denote the RAna, RAs6, and RAs3 simulations. The black and cyan lines denote the GMI-CTM and the GEOS-CTM. The other black lines denote a 2° by 2.5° simulation using the GMI-CTM with MERRA-2 assimilated fields (black thick line) as well two GMI-CTM simulations constrained by MERRA run at 2° by 2.5° (think black dashed line) and 1° by 1.25° (thin starred line). Two simulations of WACCM nudged to MERRA assimilated fields using nudging timescales of 5 hr^{-1} and 50 hr^{-1} are also shown in the thick dashed red and orange lines, respectively.

Simulation Name	Model (Reference)	Horizontal Resolution	Vertical Levels (Model Top)	Large-Scale Flow
GMI-CTM	NASA Global Modeling Initiative CTM (*Strahan et al., 2007, 2016*)	1° x 1.25°	72 (0.01 hPa)	MERRA-2 3-hourly averaged assimilated fields
GEOS-CTM	Goddard Earth Observing System Chemistry Transport Model (*Kouatchou et al., 2015*)	C90	" "	MERRA-2 3-hourly instantaneous assimilated fields
RAna	Goddard Earth Observing System, Version 5 (GEOS-5) GCM (*Suarez et al., 2008; Molod et al. 2015*)	" "	" "	MERRA-2 6-hourly analysis fields
RAs3	" "	" "	" "	MERRA-2 3-hourly averaged assimilated fields; 3 hour replay interval
RAs6	" "	" "	" "	MERRA-2 3-hourly averaged assimilated fields; 6 hour replay interval

Table 1. Details of the model integrations, where columns 3-5 correspond to the horizontal resolution, number and vertical levels and model top, and source of meteorological fields. The first simulation corresponds to an integration of the GMI-CTM at 1° by 1.25° using three hourly time-averaged assimilated fields from MERRA-2. The second simulation corresponds to an integration of the GEOS-CTM run at the cubed sphere C90 resolution (approximately 1° by 1.25°) using three hourly instantaneous assimilated fields from MERRA-2. The remaining three GEOS-5 replay simulations are run at C90 and are also constrained with MERRA-2 meteorological fields. The RAna simulation uses six-hourly instantaneous analysis fields (row 4), whereas the RAs3 and RAs6 use the time-averaged assimilated MERRA-2 fields (rows 5-6). The only difference between the RAs3 and the RAs6 simulations is the frequency at which the IAU replay cycle is performed.

Tracer (χ)	**Boundary Condition (χ_Ω)**	**Source**
Stratospheric Mean Age (Γ_{GLB})	0 over Ω	1 year/year
Periodic Oscillation (χ_{tape})	$1 + \sin(\frac{2\pi t}{1\mathrm{year}})$ within $\pm 10°$ at 100 mb	0
5-Day NH-Loss (χ_5)	1 over Ω_{MID}	$-\frac{\chi}{\tau_c}$ (τ_c = 5 days, entire atmosphere)
50-Day NH-Loss (χ_{50})	1 over Ω_{MID}	$-\frac{\chi}{\tau_c}$ (τ_c= 50 days, entire atmosphere)
Tropospheric Mean Age (Γ_{NH})	0 over Ω_{MID}	1 year/year
Stratospheric-Loss (χ_{STE})	200 ppbv above 80 hPa	$-\frac{\chi}{\tau_c}$ (τ_c = 25 days, troposphere only)

Table 2. Table of tracers, χ, integrated in the simulations. All tracers (χ) satisfy the tracer continuity equation, $(\partial_t + \mathcal{T})\chi(\mathbf{r}, t|\Omega) = S$ in the interior of the atmosphere (that is, outside of Ω), where $\mathcal{T}$ is the linear advection-diffusion transport operator and S denotes interior sources and sinks. For the first tracer Ω is taken to be all grid point in the first model level (row 2). The second tracer, χ_{tape}, is a prescribed sinusoid in a mixing ratio over 10°S-10°N at 100 mb that has a maximum at October (row 3). For three of the tropospheric tracers (rows 4-6) Ω is taken to be the NH midlatitude surface, Ω_{MID}, which is defined throughout as the first model level spanning latitudes between 30°N and 50°N. The final tracer, χ_{STE}, is set to 200 ppbv for pressures less than and equal to 80 mb, and decays uniformly in the troposphere at a loss rate $\tau_d = 25$ days $^{-1}$ (row 7).