Geophysical Research Letters

RESEARCH LETTER
10.1029/2018GL080960

Key Points:
- Different conclusions from previous studies evaluating model simulations of interhemispheric transport are reconciled
- The simulated interhemispheric exchange time τ_{ex} agrees well with observations, but the simulated SF_6 age in the SH is older than observed
- Transport from the northern extratropics into the tropics is too slow in most models, which has a larger effect on SF_6 age in the SH than τ_{ex}

Supporting Information:
- Supporting Information S1
- Supporting Information S2
- Data Set S1

Correspondence to:
H. Yang,
hyang61@jhu.edu

Citation:
Yang, H., Waugh, D.W., Orbe, C., Patra, P.K., Jöckel, P., Lamarque, J.-F., et al. (2019). Evaluating simulations of interhemispheric transport: Interhemispheric exchange time versus SF_6 age. *Geophysical Research Letters*, *46*, 1113–1120. https://doi.org/10.1029/2018GL080960

Received 17 OCT 2018
Accepted 16 JAN 2019
Accepted article online 22 JAN 2019
Published online 30 JAN 2019
Corrected 12 FEB 2019

This article was corrected on 12 FEB 2019. See the end of the full text for details.

Evaluating Simulations of Interhemispheric Transport: Interhemispheric Exchange Time Versus SF_6 Age

Huang Yang[1], Darryn W. Waugh[1,2], Clara Orbe[3], Prabir K. Patra[4], Patrick Jöckel[5], Jean-Francois Lamarque[6], Simone Tilmes[6], Douglas Kinnison[6], James W. Elkins[7], and Edward J. Dlugokencky[7]

[1]Department of Earth and Planetary Sciences, Johns Hopkins University, Baltimore, MD, USA, [2]School of Mathematics, University of New South Wales, Sydney, Australia, [3]NASA Goddard Institute for Space Studies, New York, New York, NY, USA, [4]Department of Environmental Geochemical Cycle Research, JAMSTEC, Yokohama, Japan, [5]Deutsches Zentrum für Luft- und Raumfahrt (DLR), Institut für Physik der Atmosphare, Oberpfaffenhofen, Germany, [6]National Center for Atmospheric Research (NCAR), Boulder, CO, USA, [7]Earth System Research Laboratory, NOAA, Boulder, CO, USA

Abstract Two recent studies using sulfur hexafluoride (SF_6) observations to evaluate interhemispheric transport in two different ensembles of atmospheric chemistry models reached different conclusions on model performance. We show here that the different conclusions are due to the use of different metrics and not differences in the performance of the models. For both model ensembles, the multimodel mean interhemispheric exchange time τ_{ex} agrees well with observations, but in nearly all models the SF_6 age in the southern hemisphere is older than observed. This occurs because transport from the northern extratropics into the tropics is too slow in most models, and the SF_6 age is more sensitive to this bias than τ_{ex}. Thus, simulating τ_{ex} correctly does not necessarily mean that transport from northern midlatitudes into the southern hemisphere is correct. It also suggests that more attention needs to be paid to evaluating transport from northern midlatitudes into the tropics.

Plain Language Summary Transport of air between hemispheres is important as most air pollutants are emitted predominantly in northern midlatitudes, and the rate of this interhemispheric transport influences the global distribution of these gases. Two recent studies analyzed simulations from two different model intercomparison projects and reached different conclusions on the models' ability to simulate the time scale for interhemispheric transport. In this study, we show that the different conclusions are due to the use of different metrics for quantifying the cross-equator transport time scales, rather than different performances between models. For both groups of models, the mean of the models agrees with observed interhemispheric exchange time τ_{ex} but overestimates the SF_6 age at the surface level. This occurs because τ_{ex} is an estimate of the time scale of transport across the equator, whereas the SF_6 age in the southern hemisphere quantifies transport from northern midlatitudes into the tropics as well as transport across the equator. Our results indicate that transport into the tropics is generally too slow in models, resulting in too old SF_6 ages compared to observed values, with little influence on τ_{ex}.

1. Introduction

Evaluating how well atmospheric chemical models represent the time scales for large-scale transport is a crucial aspect of evaluating the suitability of models for use in understanding past changes and for predicting future changes in chemical composition. One approach for evaluating transport in models is through comparisons of simulations with observations of trace gases that have well-known sources and sinks and whose spatial gradients are determined primarily by transport. One such trace gas is sulfur hexafluoride (SF_6), which has a very long atmospheric lifetime, a large growth rate, and only anthropogenic sources (primarily over the northern midlatitude surface). This means that SF_6 measurements can be used to evaluate transport in models (e.g., Denning et al., 1999; Gloor et al., 2007; Krol et al., 2018; Peters et al., 2004). In addition to evaluating model performance in terms of SF_6 concentrations, studies have also shown that SF_6 can be used to constrain time scales of interhemispheric transport (IHT). Historically, IHT has been quantified using an interhemispheric exchange time τ_{ex} based on a simple model in which the global atmosphere is divided into two well-mixed boxes, one for the Northern Hemisphere (NH) and the other for the

Southern Hemisphere (SH). Initial studies used surface measurements of CO_2, bomb radiocarbon, or ^{85}Kr to estimate τ_{ex} (see Czeplak & Junge, 1975, and references therein), but more recently SF_6 observations have been used (e.g., Denning et al., 1999; Geller et al., 1997; Levin & Hesshaimer, 1996; Patra et al., 2009, 2011). The exchange time τ_{ex} has been used in a wide range of atmospheric composition applications, including identifying missing sources of ozone-depleting substances (e.g., Liang et al., 2014), constraining the abundance, temporal changes, and hemispheric asymmetries in chemical loss by the hydroxyl radical (OH; e.g., Montzka et al., 2000; Liang et al., 2017; Patra et al., 2014) and constraining the sources and sinks of methane (e.g., Turner et al., 2017). More recently, SF_6 observations have been used to estimate a SF_6 age (or a_{SF6}, see Waugh et al., 2013). In contrast to τ_{ex}, for which there is a single value for a given date, SF_6 age can be calculated for any location within the atmosphere where there are SF_6 measurements. The a_{SF6} is then an estimate of the mean transit time from the NH midlatitude surface to that location and can hence provide information on the IHT if that location is chosen in the SH.

Two recent multimodel studies have used SF_6 observations to evaluate the IHT in collections of models and have drawn different conclusions of how well the models performed. First, Patra et al. (2011) compared observation-based τ_{ex} with those calculations from models in the TransCom project and showed that the multimodel mean τ_{ex} was in good agreement with the observations. In contrast, a later study by Orbe et al. (2018) examined simulations of the mean age of air from the NH midlatitude surface in models participating in the Chemistry-Climate Model Initiative (CCMI) and found simulated ages in the SH that generally exceed those inferred from SF_6 observations (i.e., the multimodel mean was older than the observed SF_6 age). This old bias in the models is not explained by differences between the ideal age and SF_6 age, suggesting CCMI models are biased old in terms of time scales for IHT.

Combining the conclusions from these two studies, one could conclude that the ensemble of TransCom models simulate realistic time scales for IHT but this transport is too slow in the CCMI ensemble. However, this would be somewhat surprising as the CCMI is a later model intercomparison with generally newer models. An alternative possibility is that the difference is due to the different metrics used and not differences in transport between model ensembles. Although τ_{ex} and SF_6 ages are both calculated from SF_6 concentrations and are similar in magnitude in certain limits, they are not identical and it is possible that models in both ensembles reproduce the observation-based τ_{ex} but exhibit old SF_6 age biases in the SH. Here we examine which of the above is true by calculating τ_{ex} and SF_6 age from models in both the TransCom and CCMI ensembles, and comparing models with ground-based observations.

2. Methods

2.1. Data

We use surface observations of SF_6 from the same surface stations as considered in Waugh et al. (2013, see supporting information Table S1) Note that station Tierra Del Fuego is renamed as Ushuaia. The measurements are from instruments operated by the Halocarbons and other Atmospheric Trace Species (HATS; HATS, 2013) and Carbon Cycle Greenhouse Gases groups at the NOAA Earth System Research Laboratory Global Monitoring Division (Dlugokencky et al., 2018). The HATS measurements include quasi-continuous measurements from in situ instruments and discrete samples collected in flasks, while Carbon Cycle Greenhouse Gases measurements are only from flask samples. For both networks, we use monthly mean values of SF_6 from files available at the respective web sites (Dlugokencky et al., 2018; HATS, 2013). Data from 1995 to 2009 are used for our calculations. This is the overlap between periods where there is data and model output.

2.2. Models

We analyze monthly mean SF_6 from models participating in the TransCom-CH4 experiment (Patra et al., 2011) and CCMI phase 1 (Morgenstern et al., 2017; Orbe et al., 2018), referred to here simply as "TransCom" and "CCMI." The TransCom and CCMI activities have very different foci: The primary focus of TransCom is simulations of carbon dioxide and methane, whereas the CCMI focuses on chemistry-climate coupling, including stratospheric ozone-climate interactions. There is no overlap between models participating in TransCom and CCMI (the individual models considered in this study are listed in Table S2), with CCMI models generally have higher spatial resolution and more vertical levels than those in TransCom.

All models participating in TransCom simulated SF_6, but only a few modeling groups in CCMI included SF_6 in their simulations. The TransCom models use, with one exception, meteorology from meteorological reanalyses (either directly in a chemical transport model or nudged within the general circulation model).

The CCMI ensemble includes models using internally generated meteorology ("free-running") and models using meteorological reanalysis fields ("specified dynamics"). We consider all models in each project and do not separate free-running or specified dynamics models. (Note, Orbe et al., 2017, 2018; and Yang, Waugh, Orbe, et al., 2018, have shown that the spread in transport among specified dynamics models is comparable to that for free-running models in CCMI.) There are small differences in the SF_6 emissions used between the TransCom and CCMI models (TransCom models used EDGAR v4.0, whereas CCMI models used EDGAR v4.2), but the τ_{ex} and SF_6 age calculations are not sensitive to these differences as both calculations are more dependent on meridional gradient of SF_6 concentrations and not on emissions.

In addition to the analysis of the above comprehensive three-dimensional chemistry-climate models, we use the idealized Advanced Global Atmospheric Gases Experiment 12-box model (Rigby et al., 2013; Yang, Waugh, & Holzer, 2018) to examine the sensitivity of the SF_6 metrics to difference aspects of the flow. The Advanced Global Atmospheric Gases Experiment model divides the atmosphere into 12 boxes that are vertically separated by stratosphere (S, <200 hPa), upper troposphere (UT, 200–500 hPa), and lower troposphere (LT, 500–1,000 hPa), while meridionally separated by NH extratropics (NE, 30–90°N), NH tropics (NT, 0–30°N), SH tropics (ST, 30–0°S), and SH extratropics (SE, 90–30°S). The transport between boxes occurs via advection and diffusion, but is dominated by the diffusive exchange (see Rigby et al., 2013). The transport parameters used are generally similar to those used in Yang, Waugh, and Holzer (2018; see Data Set S1).

SF_6 is simulated in the 12-box model by setting the concentration in the Lower Troposphere Northern Extratropics (LTNE) box equal to the monthly mean average of observations from the MHD (53.3°N) and NWR (40.0°N) stations. There is no loss in any of the tropospheric boxes, while the lifetime in stratospheric boxes is set to 850 years (Ray et al., 2017).

2.3. Metrics

We consider two metrics for time scales of transport from the NH into the SH that can be calculated from SF_6 observations: the interhemispheric exchange time τ_{ex} and the SF_6 age. We use the same methods as in Patra et al. (2011) and Waugh et al. (2013) to calculate τ_{ex} and SF_6 age, respectively, so that we can compare our results directly with these studies.

If it is assumed that there is equal air mass and nonzero SF_6 emission in each hemisphere and that total SF_6 emissions in the NH (E_n) are much larger than those in the SH (E_s), the interhemispheric exchange time τ_{ex} can be written as in (Patra et al., 2009, 2011).

$$\tau_{ex} = \left[(c_n - c_s) \left(\frac{E_n}{E_s} + 1 \right) \right] \left[\frac{E_n}{E_s} \frac{dc_s}{dt} - \frac{dc_n}{dt} \right], \quad (1)$$

where $c_{n/s}$ are the surface tracer concentrations averaged over the NH and SH. As in Patra et al. (2011), the average of the annual-mean SF_6 concentrations at BRW (71.3°N) and MLO (19.5°N) is used to estimate c_n, the average from CGO (40.7°S) and SPO (90.0°S) is used to estimate c_s, and $E_{n/s}$ are determined from EDGAR v4.2 (EDGARv4.2, 2011). The same calculation is performed using the model output, that is, the SF_6 from each model is sampled at the above four sites and used in equation (1) to derive the model-based τ_{ex}.

This calculation of τ_{ex} is an overestimate of the true interhemispheric exchange time. First, because SF_6 concentration generally increases with latitude and there are no tropical stations that are used in the calculation of the hemispheric-mean concentrations of SF_6, we overestimate c_n and underestimate c_s, which results in an overestimate of the difference $c_n - c_s$ and (from equation (1)) an overestimate of τ_{ex}. Calculations using SF_6 measurements covering a wider latitude range indicate the calculation of Patra et al. (2011) overestimates τ_{ex} by around 20%. A second error in the above τ_{ex} calculation is that the calculation should involve the difference in the hemispheric-mean mass of SF_6 throughout the troposphere and not just surface values (concentration or mass). Use of surface mean rather than tropospheric-mean values will again yield a larger hemispheric difference and an overestimated exchange times. Calculations using models indicate the tropospheric-mean values of τ_{ex} are around half those of surface-based calculations (Denning et al., 1999). Even though it is a biased estimate, we use the above calculation to be consistent with Patra et al. (2011).

Our calculation of the SF_6 age (a_{SF6}) follows the method of Waugh et al. (2013). The SF_6 age is a time lag satisfying

$$c(\mathbf{r}, t) = c_0 \left(t - a_{SF6} \right), \quad (2)$$

where $c(\mathbf{r}, t)$ is the SF_6 concentration at a location $\mathbf{r}$ and c_0 is the concentration in the source region. Ideally c_0 would be formed by averaging measurements over the source region but observations are not available to do this. Here we follow Waugh et al. (2013) and use the average SF_6 mixing ratio from the MHD, NWR, and THD stations (Table S1) to construct c_0, both for the observations and the models (in models, SF_6 outputs are firstly sampled to the above stations). The time series at the reference location $c(\mathbf{r})$ is chosen at either observational sites or sampled model grids matching the observational sites. Both $c(\mathbf{r}, t)$ and $c_0(t)$ are constructed using monthly SF_6 data or model output, but $c_0(t)$ is additionally smoothed by a 23-month window so that it is monotonically increasing. While there is only a single value of τ_{ex} (for a given date), a_{SF6} can be defined for any location in the atmosphere and is a three-dimensional field if observations or model output are available. We focus here on a_{SF6} at different surface locations.

Given the fact that τ_{ex} is a single value but SF_6 age varies spatially, there is not an exact correspondence between the two quantities. However, for a two-box model or more generally in the case when meridional gradients of surface SF_6 in each hemisphere are weak and the growth rate of SF_6 is the same throughout the troposphere, SF_6 age at SH high latitudes will be similar to τ_{ex} based on surface measurements, see section 2.4 of Waugh et al. (2013).

For both τ_{ex} and a_{SF6}, we calculate monthly mean values from January 1995 to December 2009, but we focus here on the 1995–2009 climatological annual-mean value.

3. Results

3.1. TransCom and CCMI

Using the methods described in the previous section, we calculate an interhemispheric exchange time of τ_{ex} ~1.4 years from surface observations (black circles in Figures 1a and 1b) and a surface a_{SF6} that increase from around 0.8 years at the equator to around 1.5 years at the south pole (black squares in Figures 1a and 1b). This value of τ_{ex} is consistent with previous estimates from Patra et al. (2011; and earlier studies), while a_{SF6} agrees with values reported in Waugh et al. (2013). Further, the slightly larger SF_6 age at the south pole (SP) than the interhemispheric exchange time τ_{ex} is consistent with the simple arguments in section 2.4 of Waugh et al. (2013), as $c_0 - c_{SP}$ used in the age calculation is slightly larger than $c_n - c_s$ used in the τ_{ex} calculation. Note that both the calculations of τ_{ex} and SF_6 age are based on the surface SF_6 measurements and therefore quantify the time scales of IHT from the NH surface source to the SH surface including lateral transport across the equator as well as vertical transport between boundary layer and upper free atmosphere.

The corresponding values of τ_{ex} and a_{SF6} for the TransCom and CCMI models are shown in Figures 1a and 1b, respectively. Comparison of the TransCom τ_{ex} and CCMI a_{SF6} with the corresponding observational estimates confirms the results of Patra et al. (2011) and Orbe et al. (2018), respectively: The multimodel mean τ_{ex} from the TransCom models is close to the observed τ_{ex} (Figure 1a), but the SH a_{SF6} from nearly all CCMI models exceeds the observed values (Figure 1b). However, examination of τ_{ex} from CCMI models and a_{SF6} from TransCom models shows this cannot be interpreted as more accurate transport in TransCom than CCMI models. The agreement between multimodel mean and observed τ_{ex} also holds for CCMI models (Figure 1b) while a_{SF6} for TransCom models are also biased high (Figure 1a). In other words, the difference in comparisons with observed SF_6 between TransCom and CCMI models in Patra et al. (2011) and Orbe et al. (2018) is not due to differences in the performance of the two model ensembles, but is instead due to differences in both model ensembles' ability to simulate τ_{ex} and a_{SF6}.

The differences in the models' ability to simulate τ_{ex} and a_{SF6} can also be seen in Figuress 1c and 1d, which show the relationship between a_{SF6} averaged over the southern polar cap (90–60°S) and τ_{ex} for observations and the two model ensembles. There is a high correlation between τ_{ex} and a_{SF6} from the models, that is, models with larger τ_{ex} have a larger a_{SF6}, with a similar a_{SF6}-τ_{ex} relationship for each group of models. In both observations and models, a_{SF6} is larger than τ_{ex}, but the difference is larger for models, and the observations fall below the fit to the models. Thus, models that match the observed τ_{ex} overestimate a_{SF6}, while models that match observed a_{SF6} underestimate the observed τ_{ex}.

Figure 1 answers (in the negative) the question of whether there is a difference in the TransCom and CCMI ensembles' ability to reproduce IHT time scales. However, it raises the question of why models can simulate τ_{ex} but not a_{SF6}. Although both τ_{ex} and a_{SF6} quantify IHT, they measure it differently. τ_{ex} highlights the transport across the equator from the NH to the SH, whereas a_{SF6} quantifies transport from NH midlatitudes

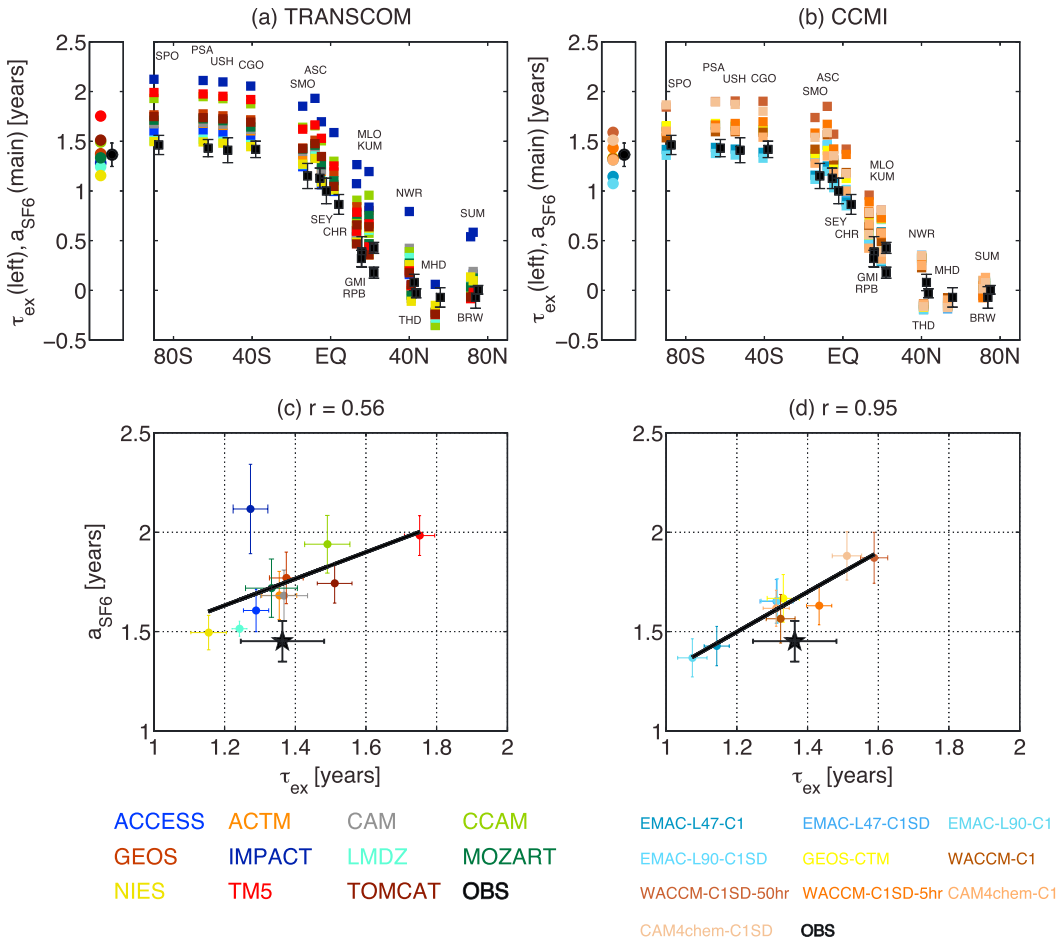

Figure 1. (a, b) Interhemispheric exchange time τ_{ex} (left plot) and SF_6 age (main plot) of models in (a) TransCom and (b) CCMI. The black circles show the observations of τ_{ex} and SF_6 age, respectively, with error bars denotes one standard deviation of interannual variability during 1995–2009, while each color symbol shows one model. (c, d) Scatterplots showing τ_{ex} versus SF_6 age over the polar cap (90–60°S average) for (c) TransCom and (d) CCMI. Colored dots are for individual models and the black stars are observations. Error bars denote one standard deviation of interannual variability. The correlation coefficients are given in the titles.

and may be sensitive to transport into the northern tropics. This means errors in models' ability to reproduce the transport into the northern tropics could have a larger impact on a_{SF6} than τ_{ex}.

An indication that errors in the transport into the northern tropics are contributing to the model a_{SF6} bias can be seen in Figure 1. The differences between simulated and observation-based a_{SF6} start to show up in the NH subtropics around 20–30°N, increase within the tropics but remain relatively unchanged in the SH. This suggests that the biased a_{SF6} in models may be related to slower transport across the NH subtropics. To test this speculation, we recalculate a_{SF6} in models and observations but using SF_6 concentrations from the two Hawaiian stations (MLO and KUM) as the reference time series (c_0 in equation (2)). These newly calculated a_{SF6} (Figure 2) show similar latitudinal variations as those shown in Figure 1, but, with the exception of one TransCom model, the newly calculated a_{SF6} are no longer biased old in the tropics and in the SH. In contrast, a_{SF6} are biased in the NH extratropics again indicating the problem lies in the simulation of NH extratropical-tropical transport. As the SF_6 concentrations in the NH tropics (specifically from the MLO station) are used to estimate c_n in τ_{ex}, any model bias in simulated northern tropical concentrations has a smaller impact on τ_{ex} (than a_{SF6}).

3.2. The 12-Box Model

To further test the influence of the NH extratropical-tropical transport on τ_{ex} and SF_6 age, we perform sensitivity simulations using the 12-box model described in section 2.2. This simple model does not capture

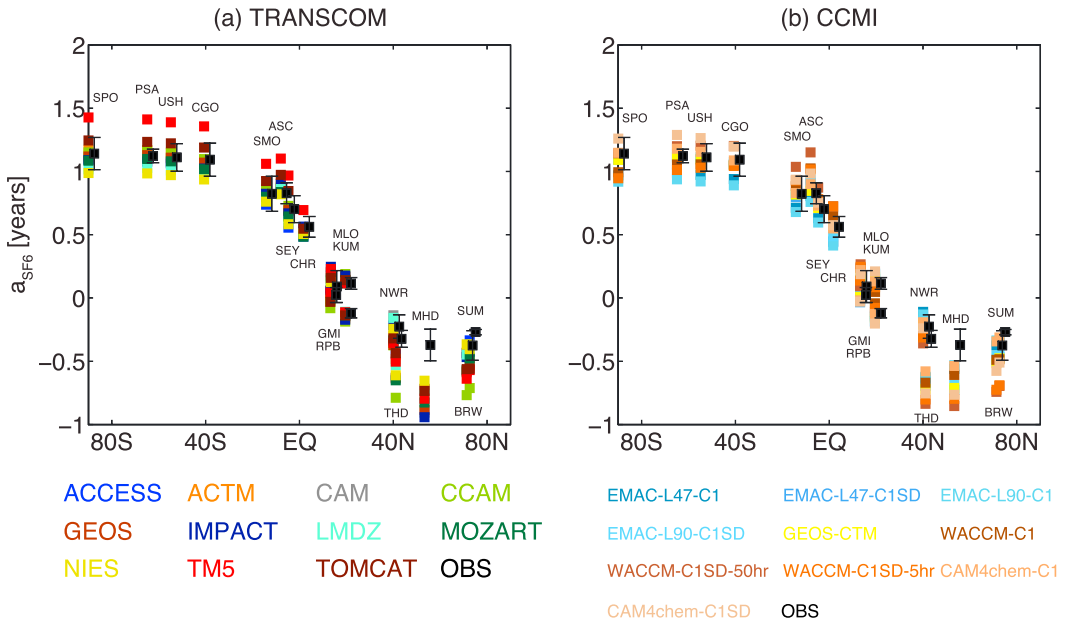

Figure 2. As in Figures 1a and 1b except the SF_6 age is calculated based on reference SF_6 time series at tropical sites (average of MLO and KUM).

all aspects of tropospheric transport, but its simplicity enables the transport between different regions to be easily perturbed, and for the impact of these changes on the two SF_6 metrics to be quantified. In the box model, the τ_{ex} calculation uses the average SF_6 concentration in the LTNE and LTNT boxes as c_n, the average of LTSE and LTST as c_s, and EDGAR v4.2 emissions (EDGARv4.2, 2011) for $E_{n/s}$. For the calculation of a_{SF6}, c_0 is the concentration in the LTNE box where the source is implemented, and a_{SF6} is calculated in all other 11 boxes but we are particularly interested in a_{SF6} in the LTSE box.

We consider five different series of perturbations (Figure 3a) where we change the diffusive time scale for (i) horizontal transport in the lower troposphere between the northern source region and northern tropics (exchange between LTNE and LTNT boxes, named as experiment S1), (ii) vertical transport between the northern lower-troposphere source region and northern upper troposphere (LTNE-UTNE, experiment S2), (iii) lateral transport in the northern upper troposphere (UTNE-UTNT, experiment UT), (iv) lateral transport within the tropical and southern lower troposphere (LTSE-LTST, LTST-LTNT, experiment LT), and (v) vertical transport in the tropics (LTST-UTST, LTNT-UTNT, experiment V). Perturbations (i)–(iii) change the transport from the northern midlatitude source region into the northern tropics, whereas (iv) and (v) change the transport in the tropics or SH.

The relationship between τ_{ex} and a_{SF6} (in the LTSE box) for each of the series of perturbations is shown in Figure 3b (different colors for each of the five types of perturbations). For the perturbations of the lateral transport into the southern tropics or vertical transport in the tropics, there is a similar sensitivity of both τ_{ex} and a_{SF6} and the values for different simulations fall along a line with slope of one. However, for perturbations of the transport from the northern extratropical source to the northern tropics, a_{SF6} is more sensitive than τ_{ex} and the simulations fall along a line with a slope larger than one, for example, for the perturbations of transport between LTNE and LTNT, a_{SF6} increases by 0.6 years over the range of the simulations, but τ_{ex} by under 0.3 years. Thus, these simple box model calculations show that a_{SF6} in the SH is more sensitive to changes in transport between the northern extratropical source and tropics than τ_{ex}. This supports our conjecture that the TransCom and CCMI models generally have too slow transport into the northern tropics, which appears as an old bias in a_{SF6} in the SH but has a smaller impact on τ_{ex}.

Perturbing only the vertical exchange between lower troposphere and upper troposphere over the northern extratropical source region (i.e., experiment S2) results in larger sensitivity of a_{SF6} than τ_{ex} as in the S1 and UT experiments. This suggests that the biased old SF_6 age among chemistry-climate models may not simply reveal slow lateral exchange between the northern extratropics and tropics but rather biases in all related

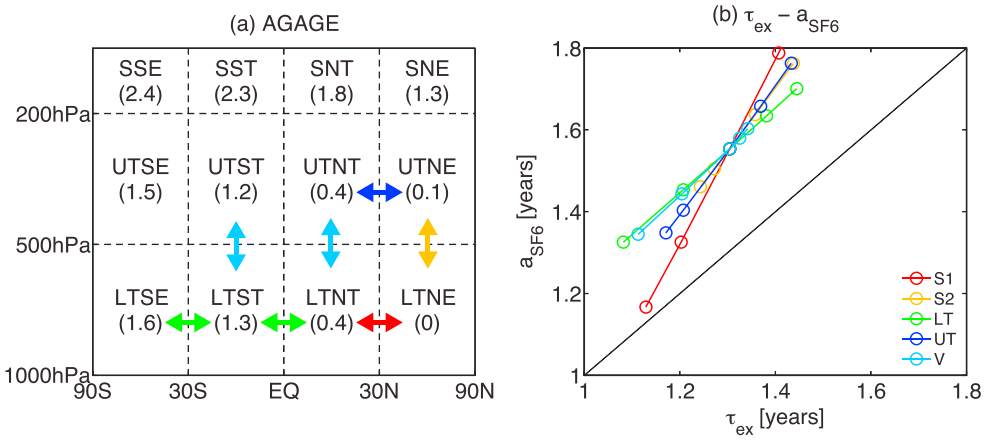

Figure 3. (a) Schematic of the AGAGE 12-box model. In the middle of each box, the name of the box is given and the corresponding climatological mean SF_6 age is additionally shown in the parenthesis. Colored arrows denote five experiments in which the diffusion coefficients between adjacent boxes are perturbed and we further examine its impact on τ_{ex} and SF_6 age in the LTSE box. (b) Variations of τ_{ex} (x axis) versus SF_6 age (y axis) in the five colored-arrows-marked experiments noted in (a). The 45° titled black line denotes the 1-1 slope that SF_6 age shows the same sensitivity to perturbation in diffusion as τ_{ex}. Therefore, a slope larger than (or smaller than) 1 denotes that, for this type of diffusion perturbation in the experiment, SF_6 age shows a larger (smaller) sensitivity than τ_{ex}. AGAGE = Advanced Global Atmospheric Gases Experiment.

Acknowledgments
We thank Scott Denning and another anonymous reviewer for constructive feedback and Steve Montzka for helpful discussions. We also thank the Centre for Environmental Data Analysis (CEDA) for hosting the CCMI data archive and the joint WCRP SPARC/IGAC Chemistry-Climate Model Initiative (CCMI) for organizing and coordinating this model data analysis activity. Huang Yang and Darryn W. Waugh acknowledge support from NSF grant AGS-1403676 and NASA grant NNX14AP58G. The EMAC simulations have been performed at the German Climate Computing Centre (DKRZ) through support from the Bundesministerium für Bildung und Forschung (BMBF). DKRZ and its scientific steering committee are gratefully acknowledged for providing the HPC and data archiving resources for the consortial project ESCiMo (Earth System Chemistry integrated Modelling). All CAM4-Chem and WACCM simulations were carried out on the Yellowstone high-performance computing platform at the National Center for Atmospheric Research. The National Center for Atmospheric Research is sponsored by the National Science Foundation. The SF_6 measurements are available from http://www.esrl.noaa.gov/gmd/hats/combined/SF6.html and ftp://aftp.cmdl.noaa.gov/data/trace_gases/sf6/flask/surface/, while the CCMI output is available from the CEDA at http://data.ceda.ac.uk//badc/wcrp-ccmi/data/CCMI-1/output/. The 12-box model parameters and results are available as Data Set S1.

transport from northern extratropical source to tropics which the "vertical trapping" over the source can also play an important role (Denning et al., 1999).

4. Conclusions

Using simulations of SF_6 from two model intercomparison projects (TransCom and CCMI) together with surface observations, we have identified the cause of different assessments of models' ability to reproduce the observed IHT in Patra et al. (2011) and Orbe et al. (2018). This is due to the different metrics (τ_{ex} or a_{SF6}) rather than the different models considered (TransCom vs CCMI). We have shown that the multimodel mean interhemispheric exchange times τ_{ex} from both model ensembles are similar to that calculated from SF_6 observations, but nearly all models overestimate the SF_6 age (a_{SF6}) in the SH. This difference in the model-data agreement occurs because the SF_6 age is more sensitive than τ_{ex} to transport from the northern extratropics into the tropics, and this transport appears be too slow in most models.

The above results suggest that caution may be required when using τ_{ex}. As we have shown, a model producing a realistic τ_{ex} does not imply that the mean time for transport from the major source of anthropogenic gases (northern midlatitudes) into the SH is simulated correctly. There may also be consequences for the common use of τ_{ex} to infer the abundance and hemispheric asymmetry in OH or sources and sinks of chemical species (such as methane or ozone-depleting substances). This potential issue requires further investigation. It may also be worth investigating the use of a three-box model (e.g., Bowman & Carrie, 2002) rather than the two-box model to study the IHT and related time scales.

Another open question, that requires further research, is what causes the consistent bias in simulated transport from northern extratropics into the tropics. Previous studies have highlighted the existence of a subtropical transport barrier (e.g., Bowman & Carrie, 2002; Bowman & Erukhimova, 2004), and the SF_6 age comparisons suggest that this barrier is too strong in the models. It is unknown whether this is due to errors in the near-surface meridional flow, convective transport, or other processes that transport air across this barrier.

References

Bowman, K. P., & Carrie, G. D. (2002). The mean-meridional transport circulation of the troposphere in an idealized GCM. *Journal of the Atmospheric Sciences*, *59*, 1502–1514.

Bowman, K. P., & Erukhimova, T. (2004). Comparison of global-scale Lagrangian transport properties of the NCEP reanalysis and CCM3. *Journal of Climate*, *17*(5), 1135–1146. https://doi.org/10.1175/1520-0442(2004)017<1135:COGLTP>2.0.CO;2

Czeplak, G., & Junge, C. (1975). Studies of interhemispheric exchange in the troposphere by a diffusion model. *Advances in Geophysics*, *18*(PB), 57–72. https://doi.org/10.1016/S0065-2687(08)60571-3

Denning, A. S., Holzer, M., Gurney, K. R., Heimann, M., Law, R. M., Rayner, P. J., & Levin, I. (1999). Three-dimensional transport and concentration of SF6a model intercomparison study (TransCom 2). *Tellus, Series B: Chemical and Physical Meteorology*, *51*(2), 266–297. https://doi.org/10.3402/tellusb.v51i2.16286

Dlugokencky, E., Lang, P., Crotwell, A., Mund, J., Crotwell, M., & Thoning, K. (2018). Atmospheric sulfur hexafluoride dry air mole fractions from the NOAA ESRL Carbon Cycle Cooperative Global Air Sampling Network, 1997-2017, Version: 2018-08-02.

EDGARv4.2 (2011). Source: European Commission, Joint Research Centre (JRC)/Netherlands Environmental Assessment Agency (PBL). Emission Database for Global Atmospheric Research (EDGAR), release version 4.2., http://edgar.jrc.ec.europa.eu

Geller, L. S., Elkins, J. W., Lobert, J. M., Clarke, A. D., Hurst, D. F., Butler, J. H., & Myers, R. C. (1997). Tropospheric SF6: Observed latitudinal distribution and trends, derived emissions and interhemispheric exchange time. *Geophysical Research Letters*, *24*(6), 675–678. https://doi.org/10.1029/97GL00523

Gloor, M., Dlugokencky, E., Brenninkmeijer, C., Horowitz, L., Hurst, D. F., Dutton, G., & Tans, P. (2007). Three-dimensional SF6 data and tropospheric transport simulations: Signals, modeling accuracy, and implications for inverse modeling. *Journal of Geophysical Research*, *112*, D15112. https://doi.org/10.1029/2006JD007973

HATS (2013). combined SF6 data set, http://www.esrl.noaa.gov/gmd/hats/combined/SF6.html

Krol, M., De Bruine, M., Killaars, L., Ouwersloot, H., Pozzer, A., Yin, Y., & Chipperfield, M. P. (2018). Age of air as a diagnostic for transport timescales in global models. *Geoscientific Model Development*, *11*(8), 3109–3130. https://doi.org/10.5194/gmd-11-3109-2018

Levin, I., & Hesshaimer, V. (1996). Refining of atmospheric transport model entries by the globally observed passive tracer distributions of 85 krypton and sulfur hexafluoride (SF6). *Journal of Geophysical Research*, *101*(D11), 16,745–16,755. https://doi.org/10.1029/96JD01058

Liang, Q., Chipperfield, M. P., Fleming, E. L., Abraham, N. L., Braesicke, P., Burkholder, J. B., & Weiss, R. F. (2017). Deriving global OH abundance and atmospheric lifetimes for long-lived gases: A search for CH3CCl3 alternatives. *Journal of Geophysical Research: Atmospheres*, *121*, 914–933. https://doi.org/10.1002/2017JD026926

Liang, Q., Newman, P. A., Daniel, J. S., Reimann, S., Hall, B. D., Dutton, G., & Kuijpers, L. J. M. (2014). Constraining the carbon tetrachloride (CCl4) budget using its global trend and inter-hemispheric gradient. *Geophysical Research Letters*, *41*, 5307–5315. https://doi.org/10.1002/2014GL060754

Montzka, S. A., Spivakovsky, H., Butler, E., Elkins, B., Lock, P., & Mondeel, J. (2000). New observational constraints for atmospheric hydroxyl on global and hemispheric scales. *Science*, *288*(5465), 500–503. https://doi.org/10.1126/science.288.5465.500

Morgenstern, O., Hegglin, M., Rozanov, E., O'Connor, F., Luke Abraham, N., Akiyoshi, H., & Zeng, G. (2017). Review of the global models used within phase 1 of the Chemistry-Climate Model Initiative (CCMI). *Geoscientific Model Development*, *10*(2), 639–671. https://doi.org/10.5194/gmd-10-639-2017

Orbe, C., Waugh, D. W., Yang, H., Lamarque, J. F., Tilmes, S., & Kinnison, D. E. (2017). Tropospheric transport differences between models using the same large-scale meteorological fields. *Geophysical Research Letters*, *44*, 1068–1078. https://doi.org/10.1002/2016GL071339

Orbe, C., Yang, H., Waugh, D. W., Zeng, G., Morgenstern, O., Kinnison, D. E., & Oman, L. D. (2018). Large-scale tropospheric transport in the Chemistry-Climate Model Initiative (CCMI) simulations. *Atmospheric Chemistry and Physics*, *18*, 7217–7235. https://doi.org/10.5194/acp-18-7217-2018

Patra, P. K., Houweling, S., Krol, M., Bousquet, P., Belikov, D., Bergmann, D., & Bian, H. (2011). TransCom model simulations of CH4 and related species: Linking transport, surface flux and chemical loss with CH4 variability in the troposphere and lower stratosphere. *Atmospheric Chemistry and Physics*, *11*, 12,813–12,837. https://doi.org/10.5194/acp-11-12813-2011

Patra, P. K., Krol, M. C., Montzka, S., Arnold, T., Atlas, E. L., Lintner, B. R., & Young, D. (2014). Observational evidence for interhemispheric hydroxyl-radical parity. *Nature*, *513*, 219–223. https://doi.org/10.1038/nature13721

Patra, P. K., Takigawa, M., Dutton, G. S., Uhse, K., Ishijima, K., Lintner, B. R., & Elkins, J. W. (2009). Transport mechanisms for synoptic, seasonal and interannual SF_6 variations and age of air in troposphere. *Atmospheric Chemistry and Physics*, *9*, 1209–1225.

Peters, W., Krol, M. C., Dlugokencky, E. J., Dentener, F. J., Bergamaschi, P., Dutton, G., & Tans, P. P. (2004). Toward regional-scale modeling using the two-way nested global model TM5: Characterization of transport using SF6. *Journal of Geophysical Research*, *109*, D19314. https://doi.org/10.1029/2004JD005020

Ray, E. A., Moore, F. L., Elkins, J. W., Rosenlof, K. H., Laube, J. C., Röckmann, T., & Andrews, A. E. (2017). Quantification of the SF_6 lifetime based on mesospheric loss measured in the stratospheric polar vortex. *Journal of Geophysical Research: Atmospheres*, *122*, 4626–4638. https://doi.org/10.1002/2016JD026198

Rigby, M., Prinn, R. G., O'Doherty, S., Montzka, S. A., McCulloch, A., Harth, C. M., & Fraser, P. J. (2013). Re-evaluation of the lifetimes of the major CFCs and CH_3CCl_3 using atmospheric trends. *Atmospheric Chemistry and Physics*, *13*(5), 2691–2702. https://doi.org/10.5194/acp-13-2691-2013

Turner, A. J., Frankenberg, C., Wennberg, P. O., & Jacob, D. J. (2017). Ambiguity in the causes for decadal trends in atmospheric methane and hydroxyl. *Proceedings of the National Acamdemy of Sciences*, *114*(21), 5367–5372. https://doi.org/10.1073/pnas.1616020114

Waugh, D. W., Crotwell, A. M., Dlugokencky, E. J., Dutton, G. S., Elkins, J. W., Hall, B. D., & Sweeney, C. (2013). Tropospheric SF_6: Age of air from the Northern Hemisphere midlatitude surface. *Journal of Geophysical Research: Atmospheres*, *118*, 11,429–11,441. https://doi.org/10.1002/jgrd.50848

Yang, H., Waugh, D. W., & Holzer, M. (2018). Decoupling the effects of transport and chemical loss on tropospheric composition: A model study of path-dependent lifetimes. *Journal of Geophysical Research: Atmospheres*, *123*, 2320–2335. https://doi.org/10.1002/2017JD027871

Yang, H., Waugh, D. W., Orbe, C., Zeng, G., Morgenstern, O., Kinnison, D. E., & Schofield, R. (2018). Large-scale transport into the Arctic: The roles of the midlatitude jet and the Hadley Cell. *Atmospheric Chemistry and Physics Discussions*. https://doi.org/10.5194/acp-2018-841

Erratum

In the originally published version of this article, author Patrick Jöckel's name was incorrectly published as Patrick Jöeckel. This error has since been corrected, and this version may be considered the authoritative version of record.