Investigating stratospheric changes between 2009 and 2018 with halogenated trace gas data from aircraft, AirCores, and a global model focusing on CFC-11

Johannes C. Laube\textsuperscript{1,2}, Emma C. Leedham Elvidge\textsuperscript{2,3}, Karina E. Adcock\textsuperscript{2}, Bianca Baier\textsuperscript{4,5}, Carl A. M. Brenninkmeijer\textsuperscript{6}, Huilin Chen\textsuperscript{7}, Elise S. Droste\textsuperscript{2}, Jens-Uwe Grooß\textsuperscript{1}, Pauli Heikkinen\textsuperscript{8}, Andrew J. Hind\textsuperscript{2}, Rigel Kivi\textsuperscript{8}, Alexander Lojko\textsuperscript{2,9}, Stephen A. Montzka\textsuperscript{5}, David E. Oram\textsuperscript{2}, Steve Randall\textsuperscript{10}, Thomas Röckmann\textsuperscript{11}, William T. Sturges\textsuperscript{2}, Colm Sweeney\textsuperscript{4}, Max Thomas\textsuperscript{2}, Elinor Tuffnell\textsuperscript{2}, and Felix Ploeger\textsuperscript{1,12}

\textsuperscript{1}Institute of Energy and Climate Research: Stratosphere, Jülich Research Centre, Jülich, 52428, Germany
\textsuperscript{2}School of Environmental Sciences, University of East Anglia, Norwich, NR4 7TJ, United Kingdom
\textsuperscript{3}Faculty of Science, University of East Anglia, Norwich Research Park, Norwich, NR4 7TJ, United Kingdom
\textsuperscript{4}Cooperative Institute for Research in Environmental Sciences, University of Colorado Boulder, Boulder, CO 80309, USA
\textsuperscript{5}Global Monitoring Division, National Oceanic and Atmospheric Administration, Boulder, CO 80305-3337, USA
\textsuperscript{6}Air Chemistry Division, Max Planck Institute for Chemistry, Mainz, 55128, Germany
\textsuperscript{7}Centre for Isotope Research, University of Groningen, Groningen, 9747 AG, the Netherlands
\textsuperscript{8}Space and Earth Observation Centre, Finnish Meteorological Institute, Sodankylä, 99600, Finland
\textsuperscript{9}Department of Climate and Space Sciences and Engineering, University of Michigan, Ann Arbor, MI 48109-2143, USA
\textsuperscript{10}Random Engineering Ltd., Felixstowe, IP11 9SL, United Kingdom
\textsuperscript{11}Institute for Marine and Atmospheric Research Utrecht, Utrecht University, Utrecht, 3508 TA, the Netherlands
\textsuperscript{12}Institute for Atmospheric and Environmental Research, University of Wuppertal, 42119 Wuppertal, Germany

Correspondence: Johannes C. Laube (j.laube@fz-juelich.de)

Received: 22 January 2020 – Discussion started: 14 February 2020
Revised: 13 June 2020 – Accepted: 16 July 2020 – Published: 20 August 2020

Abstract. We present new observations of trace gases in the stratosphere based on a cost-effective sampling technique that can access much higher altitudes than aircraft. The further development of this method now provides detection of species with abundances in the parts per trillion (ppt) range and below. We obtain mixing ratios for six gases (CFC-11, CFC-12, HCFC-22, H-1211, H-1301, and SF\textsubscript{6}), all of which are important for understanding stratospheric ozone depletion and circulation. After demonstrating the quality of the data through comparisons with ground-based records and aircraft-based observations, we combine them with the latter to demonstrate its potential. We first compare the data with results from a global model driven by three widely used meteorological reanalyses. Secondly, we focus on CFC-11 as recent evidence has indicated renewed atmospheric emissions of that species relevant on a global scale. Because the stratosphere represents the main sink region for CFC-11, potential changes in stratospheric circulation and troposphere–stratosphere exchange fluxes have been identified as the largest source of uncertainty for the accurate quantification of such emissions. Our observations span over a decade (up until 2018) and therefore cover the period of the slowdown of CFC-11 global mixing ratio decreases measured at the Earth’s surface. The spatial and temporal coverage of the observations is insufficient for a global quantitative analysis, but we do find some trends that are in contrast with expectations, indicating that the stratosphere may have contributed to the slower concentration decline in recent years.

Further investigating the reanalysis-driven model data, we find that the dynamical changes in the stratosphere required to explain the apparent change in tropospheric CFC-11 emissions after 2013 are possible but with a very high uncertainty range. This is partly caused by the high variability of mass flux from the stratosphere to the troposphere, especially at...
timescales of a few years, and partly by large differences between runs driven by different reanalysis products, none of which agree with our observations well enough for such a quantitative analysis.

1 Introduction

Many halogenated trace gases are strong greenhouse gases and/or are involved in the ongoing depletion of the ozone layer; therefore, observations of these in the stratosphere are valuable. Moreover, measurements of some of these species allow us to constrain changes in stratospheric circulation and transport across the tropopause. An analytical challenge is posed by the low abundances of many such gases, in combination with the low ambient pressures found in this region of the atmosphere. Another challenge is the ability to reach the stratosphere as even the highest-flying research aircraft can only reach altitudes just above 20 km (Schauffler et al., 2003; von Hobe et al., 2013). This is modest considering that the stratosphere extends to around 50 km. Large high-altitude balloons can reach altitudes of up to about 36 km (Engel et al., 2009; Ray et al., 2017), but due to the heavy payloads, they are increasingly difficult to fly given the risks for people living in landing areas and the cost or risk from lifting gases such as helium or hydrogen. Satellite (or aircraft) remote sensing plays an important role and can offer a global picture for some gases (Stiller et al., 2008; Santee et al., 2013; Harrison et al., 2019), but measurement precision and altitude resolution are often limited. They are also indirect observations and require continued validation through independent in situ methods. Generally, the mentioned platforms are rather expensive, ranging from costs of the order of EUR 10 000 per flight hour for aircraft to satellite costs of millions of euros. The relatively recently developed AirCore technique (Karion et al., 2010), with flight costs of below EUR 2000 (depending on the setup), offers a cost-effective alternative. AirCores, which were named due to similarities to ice cores extracted from glaciers, are based on the concept of flying a very long lightweight coiled piece of stainless-steel tubing on a weather balloon. The tube is open on one end and therefore empties naturally upon ascent as ambient pressures decrease. During descent a full vertical profile of air is collected between the balloon’s burst altitude (up to 36 km) and ground level. This technology has been demonstrated before but for providing measurements of only a handful of higher abundance trace gases such as CO\textsubscript{2} and CH\textsubscript{4} (Karion et al., 2010; Membrive et al., 2017; Engel et al., 2017) and their isotopic composition (Mrozek et al., 2016; Paul et al., 2016).

However, due to the limited amount of air that is captured by AirCores, no ozone-depleting substances (ODSs) have been investigated yet, as their abundances are well below 1 ppb (parts per billion). The importance of such observations is, however, demonstrated by the following example. The recent work by Montzka et al. (2018) on renewed emissions of CFC-11 has received much attention since it indicates a substantial and ongoing breach of the global treaty designed to prevent the destruction of the stratospheric ozone layer: the Montreal Protocol on Substances that Deplete the Ozone Layer. According to their study, global CFC-11 emissions increased by 13 ± 5 Gg yr\textsuperscript{−1} when comparing periods before and after 2012 with the caveat that up to 50 % of that derived emission change might actually be attributable to changes in stratospheric processes or dynamics. More recently, Rigby et al. (2019) found similar global increases of 11–17 Gg yr\textsuperscript{−1} over 2014–2017 vs. the 2008–2012 average, and they also pinpointed a concurrent emissions increase source of 7.0±3.0 Gg yr\textsuperscript{−1} to eastern mainland China. However, they found no emission increases in other parts of the world covered by regular ground-based observations. This could mean that some of these emission increases have arisen in regions where no such measurements are available. An alternative explanation, i.e. the possibility of a sustained change to the amount of CFC-11 exchanged between the troposphere and the stratosphere as the driving mechanism for at least a part of the anomaly, has, however, not been ruled out so far.

2 Methods

Dry air mole fractions of halogenated trace gases were derived from air samples collected on board three different platforms: a passenger aircraft (CARIBIC; Brenninkmeijer et al., 2007) flying at altitudes of 8–13 km (11 flights, 2009–2016), a research aircraft (Leedham Elvidge et al., 2018) accessing higher altitudes of 9–21 km (M55 Geophysika, five campaigns, 2009–2017), and the first measurements of such gases with the relatively recently developed AirCore methodology (Karion et al., 2010; 8–30 km, 15 flights in Finland and the UK, 2016–2018). The aircraft data have partly been published before (Leedham Elvidge et al., 2018; Laube et al., 2013). The balloon-based AirCore technique was developed further, mainly through the use of specially designed tubing that maximises the amounts of air collected in the stratosphere, as well as through a novel subsampling technique that minimises the use of contamination-prone materials. The amount of retrievable stratospheric air, however, is still more than 2 orders of magnitude smaller than from aircraft-based sampling techniques. With laboratory analytical improvements compensating for this, the AirCore measurements show good precisions (ranging from 0.2 % to 3.3 % compared with 0.4 % to 1.1 % for aircraft samples) and excellent agreement with the aircraft data. The other important challenge for AirCore measurements of halocarbons is to ensure that the air is not contaminated throughout the entire sampling and subsampling process. Contaminations can arise from leakages and/or halocarbon-emitting materials (such as organic polymers) in the AirCore itself, in the CO\textsubscript{2} analyser.
Figure 1. Aircraft- and balloon-based mixing ratios of six halogenated trace gases in the upper troposphere and stratosphere as compared to the NOAA/GMD ground-based northern hemispheric GGGRN time series (https://www.esrl.noaa.gov/gmd/, last access: 11 January 2020). HCFC-22 has a significant sink process in the troposphere and therefore exhibits stronger inner-hemispheric gradients. To illustrate that, we compare the mid-latitude station at Mace Head, Ireland, with the subtropical station at Mauna Loa, Hawaii. Lower mixing ratios generally represent higher altitudes. For all gases except SF$_6$, some higher-altitude data are not shown to better demonstrate the good comparability of near-tropopause data to the NOAA time series. The complete corresponding data including uncertainties can be found in the Supplement (see also Figs. S1 to S4).

system including the pump, or in the subsampling system. Importantly, for all compounds reported here, mixing ratios in the stratosphere are much lower than in even remote tropospheric regions, let alone near sources of these gases. In addition, almost all of the contamination possibilities would affect the entire profile as an AirCore is essentially one air sample. This would become apparent in the correlations of the species with each other, which are very compact in the stratosphere. In the absence of such correlation breakdowns (see Figs. 1, 2, and S1 to S4), we therefore conclude that such contaminations are at undetectable levels in the dataset presented here. More details can be found in Table 1 and the Supplement.

All samples were processed with a previously described analytical system and methodology (Laube et al., 2010, 2012) using cryogenic extraction and pre-concentration, followed by gas chromatographic separation and detection with a high-sensitivity mass spectrometer. Trace gas measurements from this system as well as mean ages of air (AoAs, i.e. average stratospheric transit times; see section 3.1 for more details) calculated from these have been shown to compare very well with those of other internationally recognised measurements over several decades (Leedham Elvidge et al., 2018; Laube et al., 2013; Trudinger et al., 2016).

Stratospheric trends at AoA surfaces were derived by fitting second- and third-order polynomials (depending on whether an inflexion point was observed) to the respective correlations of mixing ratios and AoAs. The formulas of the polynomials were then used to interpolate onto the AoA surfaces (1, 2, 3, and/or 4 years, depending on which AoA range was covered) for each flight. To test the uncertainty of this method, the data for each flight were first replicated four times, where each replicate was modified by plus or minus the uncertainty in the mixing ratio and mean age uncertainties. This resulted, for each data point, in the average plus minimum and maximum value for both mixing ratio and AoA. Subsequently, $5n$ ($n$ being the number of data points available for each flight) random samples were drawn (repeat draws possible) with a bootstrap algorithm (as in Volk et al., 1997; Laube et al., 2013), and a second- or third-order poly-
nominal again fitted. This procedure was repeated 500 times for each flight, resulting in an average mixing ratio and an uncertainty range at each AoA surface. The derived mixing ratios were subsequently used to produce linear regressions over time, including a weighting by the inverse uncertainties of the individual CFC (chlorofluorocarbon) mixing ratios. The bootstrapping algorithm (500 repeat draws, repeat draws possible) was used again to ensure that the derived slope uncertainties were not underestimated and that individual high or low points did not bias the slope estimates.

Observation-based data were compared to model output from the Chemical Lagrangian Model of the Stratosphere (CLaMS), a Lagrangian chemical transport model with advective transport calculated from three-dimensional forward trajectories and an additional parameterisation for small-scale turbulent mixing (McKenna et al., 2002). Potential temperature is used as vertical coordinate throughout the stratosphere with vertical velocity estimated from the total diabatic heating rate. Further model details and the chemistry scheme used are described in Pommrich et al. (2014). For the simulations used in this study, CLaMS was driven with horizontal winds and diabatic heating rates from three alternative meteorological reanalysis datasets: ERA-Interim (from the European Centre for Medium-Range Weather Forecasts, ECMWF), JRA-55 (from the Japan Meteorological Agency), and MERRA-2 (from NASA). For more information on methods, calibrations, and modelling, as well as additional data, please see the Supplement.

3 Results and discussion

3.1 Observational data overview and comparisons

Our data are based on measurements of air samples collected in the upper troposphere and stratosphere of the northern hemisphere using aircraft and weather balloons between 2009 and 2018. Figure 1 shows the obtained mixing ratios alongside the northern hemispheric “background” time series derived through the combination of observations at various ground-based stations within the National Oceanic and Atmospheric Administration Global Monitoring Division’s Global Greenhouse Gas Reference Network (NOAA/GMD GGGRN). It is apparent that both the aircraft and the balloon data follow the ground-based trends quite well for all six gases. Slightly enhanced mixing ratios can often be observed in the vicinity of the tropopause (see also Figs. S5 and S6), mostly due to recent influences from regional emissions (Kloss et al., 2014; Leedham Elvidge et al., 2015; Oram et al., 2017). This is especially pronounced in the research aircraft data from 2017, which belong to a campaign (Höpfner et al., 2019) exploring the atmospheric composition above the polluted Asian monsoon region (Randel et al., 2010; Vogel et al., 2019). It is, however, worth noting that most species’ enhancements are not significantly higher than the combined measurement uncertainties, which demonstrates the importance of the consistency of the datasets and therefore the quality of the stratospheric record. Figure 1 also illustrates the much improved temporal density that AirCore observations have provided from 2016 onwards (in comparison to aircraft campaigns), especially at altitudes above 15 km, which are out of the reach of all but a few research aircraft.

In the stratosphere, trace gases typically exhibit compact interspecies correlations (Schaufler et al., 2003; Volk et al., 1997), and some gases (such as SF$_6$) can be utilised to de-
Figure 3. Panel (a) shows a comparison of observation-based CFC-11 mixing ratio trends at mean ages of air of 2 (blue), and 4 (red) years with those from the CLaMS model run driven by the ERA-Interim reanalysis (grey and yellow) in the northern hemispheric stratosphere. The latter have been derived as averages between 30 and 90° N. The dashed and dotted lines correspond to regression lines (weighted by their 1σ standard error for observations) and an illustration of their 2σ uncertainties over the time periods displayed. Panel (b) shows the same comparison but at mean ages of 1 and 3 years. The numerical values can be found in Table 2.

3.2 Comparisons with model data using different reanalyses

We first focus on a comparison of model simulations with the aircraft and AirCore data. Also shown in Fig. 2 are data from simulations with the Chemical Lagrangian Model of the Stratosphere (CLaMS; McKenna et al., 2002; Pommrich et al., 2014). The latter was driven alternatively by three commonly used meteorological reanalyses, i.e. ERA-Interim, JRA-55, and MERRA-2 (Dee et al., 2011; Kobayashi et al., 2015; Gelaro et al., 2017). These newest available meteorological reanalysis datasets provide the best guess of the current state of the atmosphere. We use the differences between to quantify the uncertainty in our knowledge of the stratospheric circulation and its changes. The model was sampled at coordinates and times coinciding with those of the observations. Results from all three runs are similar to those from observations in the case of the correlation of CFC-11 with CFC-12. The CFC-11–AoA correlation in Fig. 2 is a measure of the speed of the main stratospheric overturning circulation as it reflects, in an integrated way, the speed and pathway of trace gas transport through the stratosphere. Here, the model data for both ERA-Interim and JRA-55 remain close to the observed values throughout the range. The MERRA-2-based data does, however, stand out producing higher AoAs at similar stratospheric CFC-11 mixing ratios and an increasing discrepancy with increasing AoA. As noted by Ploeger et al. (2019), the MERRA-2 reanalysis has a slower stratospheric circulation, and our observational evidence strongly indicates that it is indeed too slow. This is a consistent feature, which is also apparent when comparing with MERRA-2-based data from before 2016 (not shown in Fig. 2). The details of the causing mechanisms could be complex and are beyond the scope of this work.

3.3 Long-term trends of trace gases in the stratosphere

Focusing on the details of the correlations in Fig. 2, we investigate whether there are indications here that might partly explain the recent changes in the tropospheric trend of CFC-11. Most air enters the stratosphere in the tropics and is then transported poleward. CFC-11 and CFC-12 molecules are mostly destroyed in the tropical stratosphere (Douglass et al., 2008). Transport of the remainder of these gases to the...
Table 1. Comparison of average measurement uncertainties (derived as the average of 1 standard deviation from repeated working standard or air sample measurements) of the research aircraft campaign in 2016; all AirCore flights; and some AirCore sample repeats for CFC-11, CFC-12, H-1211, H-1301, HCFC-22, and SF6. For AirCore uncertainties, the average working standard uncertainty over 3 years was used as it is (a) more representative of the entire measurement period and (b) generally comparable or worse than precisions derived from sample repeats. AirCore-based precisions are generally slightly worse than those achieved with the larger aircraft-based samples but still much smaller than mixing ratio gradients observed in the stratosphere.

<table>
<thead>
<tr>
<th>Trace gas</th>
<th>Average precision (%) of aircraft 2016 measurements</th>
<th>Average precision (%) of AirCore 2016–2018 standards</th>
<th>Average precision (%) of AirCore 2017 sample repeats</th>
</tr>
</thead>
<tbody>
<tr>
<td>CFC-11 (CFCl3)</td>
<td>0.4</td>
<td>0.9</td>
<td>1.2</td>
</tr>
<tr>
<td>CFC-12 (CF2Cl2)</td>
<td>1.1</td>
<td>1.2</td>
<td>0.8</td>
</tr>
<tr>
<td>H-1211 (CF2ClBr)</td>
<td>0.6</td>
<td>1.9</td>
<td>1.0</td>
</tr>
<tr>
<td>H-1301 (CF3Br)</td>
<td>0.6</td>
<td>3.3</td>
<td>2.3</td>
</tr>
<tr>
<td>HCFC-22 (CHF2Cl)</td>
<td>0.6</td>
<td>0.9</td>
<td>0.2</td>
</tr>
<tr>
<td>SF6</td>
<td>0.4</td>
<td>0.9</td>
<td>0.6</td>
</tr>
</tbody>
</table>

For the other three gases that have sufficient measurement precisions for such an analysis (i.e. CFC-12, H-1211, and HCFC-22), we also find a picture that does not agree well with their tropospheric trends (Table 2). Both CFC-12 and H-1211 have been in decline in the troposphere since the mid-2000s and decreased by ∼6% and ∼20% between late 2009 and late 2018, respectively (Fig. 1), whereas tropospheric HCFC-22 mixing ratios have increased monotonically (and by ∼25% during our observation period) since the trace gas appeared in the atmosphere several decades ago, albeit with a recent slowdown. In contrast, in the stratosphere, we find that CFC-12 decreased at all mean age surfaces but not as much as in the troposphere (∼0.9% to −4%); HCFC-22 increased disproportionally by 30% to 38%; and H-1211 decreased, but only at a mean age of 1 year (−9%). No significant change occurred at 2 years, and 9% to 22% increases were observed at 3 and 4 years. For the latter three gases, this unexpected behaviour could be partly related to changes in tropospheric trends in the period leading up to 2009, as a significant part of the air at certain mean age levels is much older than the mean age itself. However, these effects should subside over the decade that our observations span, especially for H-1211, which is the shortest-lived gas of the four. In addition, CFC-11 should not be affected as it has been decreasing for much longer. The underlying mechanisms are likely complex.

The only straightforward possibility to generate positive CFC-11 trends in the stratosphere between 2009 and 2018 would be an increase in the air fractions that have younger and older residence times than the inferred mean age. Such a 2-fold increase would maintain the same AoA, but would influence the mixing ratios observed at the AoA surfaces in different ways. If the increased older air fraction had been in the stratosphere for long enough, it would have already influence the mixing ratios observed at the AoA surfaces in different ways. If the increased older air fraction had been in the stratosphere for long enough, it would have already
Table 2. Temporal trends and their 2σ uncertainties of CFC-11, CFC-12, HCFC-22, and H-1211 mixing ratios at AoAs of 1, 2, 3, and 4 years. These slopes correspond to an uncertainty-weighted regression line fitted to the data in Figs. 3, and S8–S12, with two exceptions: (1) the data from the Asian monsoon campaign in 2017 was excluded as this region is not representative of northern hemispheric stratospheric air and (2) all data at mean ages above 3.5 years from winter campaigns in high latitudes were also excluded as they might contain polar vortex air, which is equally unrepresentative. Model-based slopes were derived over the same period as observational data (August 2009–August 2018), except for JRA-55 and MERRA-2, where data were only available until the end of 2017.

<table>
<thead>
<tr>
<th></th>
<th>Slope obs. (ppt yr⁻¹)</th>
<th>Uncertainty (ppt yr⁻¹)</th>
<th>Trend (% per decade)</th>
<th>Slope (ppt yr⁻¹)</th>
<th>Uncertainty (ppt yr⁻¹)</th>
<th>Trend (% per decade)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CFC-11</td>
<td>0.69</td>
<td>1.77</td>
<td>1.25</td>
<td>0.59</td>
<td>1.54</td>
<td>1.81</td>
</tr>
<tr>
<td></td>
<td>1.54</td>
<td>1.60</td>
<td>2.12</td>
<td>1.74</td>
<td>2.2</td>
<td>2.52</td>
</tr>
<tr>
<td></td>
<td>0.22</td>
<td>0.47</td>
<td>0.61</td>
<td>0.20</td>
<td>0.27</td>
<td>0.62</td>
</tr>
<tr>
<td></td>
<td>0.18</td>
<td>0.23</td>
<td>0.30</td>
<td>0.18</td>
<td>0.23</td>
<td>0.30</td>
</tr>
<tr>
<td>CFC-12</td>
<td>-1.96</td>
<td>0.45</td>
<td>-0.38</td>
<td>-1.33</td>
<td>1.90</td>
<td>2.20</td>
</tr>
<tr>
<td></td>
<td>1.90</td>
<td>2.52</td>
<td>3.52</td>
<td>5.36</td>
<td>3.09</td>
<td>2.52</td>
</tr>
<tr>
<td></td>
<td>0.37</td>
<td>0.62</td>
<td>1.20</td>
<td>0.39</td>
<td>0.60</td>
<td>1.08</td>
</tr>
<tr>
<td></td>
<td>0.28</td>
<td>0.27</td>
<td>0.62</td>
<td>0.27</td>
<td>0.27</td>
<td>0.62</td>
</tr>
<tr>
<td></td>
<td>0.24</td>
<td>0.39</td>
<td>0.76</td>
<td>0.39</td>
<td>0.51</td>
<td>0.76</td>
</tr>
<tr>
<td>HCFC-22</td>
<td>6.15</td>
<td>6.16</td>
<td>5.98</td>
<td>5.67</td>
<td>0.15</td>
<td>0.14</td>
</tr>
<tr>
<td></td>
<td>0.15</td>
<td>0.14</td>
<td>0.18</td>
<td>0.18</td>
<td>30.5</td>
<td>33.4</td>
</tr>
<tr>
<td></td>
<td>30.5</td>
<td>33.4</td>
<td>36.0</td>
<td>38.2</td>
<td>30.5</td>
<td>33.4</td>
</tr>
<tr>
<td>H-1211</td>
<td>-0.031</td>
<td>0.000</td>
<td>0.013</td>
<td>0.013</td>
<td>0.008</td>
<td>0.007</td>
</tr>
<tr>
<td></td>
<td>0.008</td>
<td>0.007</td>
<td>0.009</td>
<td>0.009</td>
<td>-9.0</td>
<td>0.2</td>
</tr>
<tr>
<td></td>
<td>-9.0</td>
<td>9.1</td>
<td>22.4</td>
<td>22.4</td>
<td>-9.0</td>
<td>9.1</td>
</tr>
</tbody>
</table>

The nearer mentioned possibility to at least partly explain such trends could include an acceleration of air mass transport through the lower tropical stratosphere (i.e., below the main sink region of CFC-11) as, for example, CLaMS–ERA-Interim qualitatively shows over the relevant period (Fig. S15). However, when compared with ERA-Interim-based model data at the same transport times (Fig. 3), the model results show a different CFC-11 trend in the lower stratosphere. In fact, the model- and observation-based trends at mean ages of 1 and 2 years do not agree within 2 standard deviations. This discrepancy is likely related to a known problem with ERA-Interim, which generally overestimates the speed of the circulation in that lower stratospheric region (Dee et al., 2011; Ploeger et al., 2012). At larger mean ages, we find better agreement between the observations and the model with the model data even reproducing the observed insignificant trend. JRA-55-based model trends are very similar to those from the ERA-Interim-based analysis, whereas the MERRA-2 reanalysis shows larger differences to observations, both in terms of mixing ratios and trends (Table 2, Figs. 3, and S8–S12). The generally limited comparability of model and observations sheds some light on the ability of current reanalysis products to quantify structural changes in stratospheric circulation patterns.

3.4 Mass flux estimates of CFC-11

Nevertheless, we use the reanalysis-driven model data as the best available means to derive the downward mass flux of CFC-11 through the extratropical tropopause, i.e., the quantity describing how much CFC-11 is transported back to the troposphere. Comparing the three simulations driven with three different reanalyses provides an estimate of uncertainty due to representations of stratospheric circulation changes. A temporal increase of the stratosphere-to-troposphere mass flux could cause changes to the tropospheric trend of CFC-11, which would look like renewed emissions. Such a flux increase could be consistent with the observed increases in CFC-11 mixing ratios on AoA surfaces (Sect. 3.3) if accompanied by an increased fraction of air entering the stratosphere without passing through the main CFC-11 sink region in the lower tropical stratosphere (and instead entering, for example, through the Asian summer monsoon). The NOAA/GMD tropospheric time series of CFC-11 serves as the boundary condition for the model, and consequently in the absence of stratospheric changes, the temporal trend of the mass flux should be similarly negative and of a similar magnitude. The model generally reflects this reason-
ing over longer time periods as can be seen in Fig. 4. We then follow the approach by Montzka et al. (2018) to investigate whether the changes to the tropospheric trend around 2013 might partly be caused by more CFC-11 being transported back into the troposphere. For that purpose, we split the data into two periods: before and after 2013. Independent of which definition of the tropopause is being used (see the Supplement for details), we find an increase in the mass flux of around 37 Gg yr$^{-1}$ after 2013 for CLaMS–ERA-Interim. This would explain 270% of the observed slowdown of CFC-11 mixing ratio decreases after 2013 when comparing to the $13 \pm 5$ Gg yr$^{-1}$ emission increase inferred by Montzka et al. (2018). At first glance, this very high stratospheric contribution is not consistent with the findings of both Montzka et al. (2018) and Rigby et al. (2019), who estimated 40% to 60% of the slowdown to belong to renewed emissions. However, the global stratosphere-to-troposphere mass flux is very large compared to the amount of unexplained emissions, and a direct quantitative comparison is not possible, as explained in the following. When repeating the same model run, but with an artificial tropospheric CFC-11 trend that continues to decrease linearly after 2013 the mass flux remains very similar to the reference simulation (difference of $<0.6$ Gg yr$^{-1}$). This translates into a minor influence of recent tropospheric trend changes on these stratospheric fluxes and therefore confirming that this signal is indeed driven by stratospheric changes in the ERA-Interim world. However, this pronounced turnaround in 2013 is not a consistent feature for all three reanalyses, as the JRA-55 run, despite producing such a similar picture in the correlation comparisons (Fig. 2), in fact shows a further decrease of 0.4 Gg yr$^{-1}$ (equivalent to $-3\%$ of the new emissions signal) after 2013. The main reason for that discrepancy is that, as opposed to ERA-Interim, JRA-55 does not show a substantial change to the mass flux around 2013. Coming back to the pre- and post-2013 analysis, CLaMS–MERRA-2 results are in between the other two with 18.2 Gg yr$^{-1}$ (135%), but have the least credibility as demonstrated by the poor comparability with observations. The main issue connected with such an analysis is illustrated in Fig. 4. With annual changes of up to 21%, the variability of the CFC-11 mass flux from the stratosphere to the troposphere is an order of magnitude higher than the 2013 change of 2% to 5% that we are trying to quantify. Some of that mass flux variability occurs over several years, which severely limits the capability of quantitatively determining trend changes between an 11- and a 5-year period. It should, however, be re-emphasised that a mass flux trend analysis over longer periods would be expected to work better and this is indeed what we find for ERA-Interim and JRA-55. Between 2002 and 2017 the CFC-11 flux from a linear regression of the model output driven by these two reanalyses decreases by 10.5% and 13.1%, respectively, which is comparable to the $\sim 11\%$ tropospheric decrease over the same period. MERRA-2 again produces an outlier with only a 3.2% decrease during those 16 years. The recent findings by Ray et al. (2020) of the QBO (Quasi-biennial Oscillation) significantly modulating the variability of long-lived trace gases at the surface are qualitatively consistent with our findings for both shorter and longer periods. However, a quantification of this modulation is currently limited by the uncertainties connected to the meteorological reanalyses in the stratosphere. As shown in Fig. 4, the mass fluxes from the three CLaMS-reanalysis runs show some covariation on QBO timescales but at the same time also some significant differences which include offsets, long-term trends, the magnitude of the variations, and the timing of changes.

4 Conclusions

To summarise, we present new observations of six halogenated trace gases in the stratosphere obtained from applying a further-developed AirCore technology. These ob-

Figure 4. The annually averaged stratosphere-to-troposphere mass flux of CFC-11 through the tropopause between 2002 and 2018 for CLaMS model runs driven by MERRA-2 (green), JRA-55 (blue) and ERA-Interim (black) reanalyses including a linear regression for the period until 2013 (dashed). The red line originates from an ERA-Interim sensitivity run for which tropospheric CFC-11 was forced to continue to decrease at the same rate as before 2013. Shown in grey and on the right-hand y axis are the two corresponding time series of tropospheric CFC-11 mixing ratios (i.e. the real one, solid, and the one with the forced decrease, dashed). The annual average has been calculated by applying a 12-month running mean to the time series.
servations are consistent with ground-based measurements of the same species at remote locations. They compare well to aircraft-based observations, have good precisions, and offer a viable low-cost method for directly observing ozone-depleting gases and circulation tracers in the stratosphere at enhanced temporal and spatial resolutions. The derived mixing ratios and mean stratospheric residence times, both from aircraft and AirCore data, enable the assessment of the performance of the three most modern currently available meteorological reanalysis packages. The ERA-Interim- and JRA-55-derived model data compare better, whereas the MERRA-2-based data exhibit distinctly slower transport through most of the region covered here.

From a further analysis of the observational data at certain stratospheric transport times, we also find insignificant to positive trends (within 1 standard deviation) of CFC-11 mixing ratios in the lower stratosphere between 2009 and 2018 ranging from 3% to 10%. This is surprising and in contrast to expectation from the tropospheric abundances, which have been decreasing by about 6% over that period. Similarly derived trends for CFC-12, HCFC-22, and H-1211 are also not in good agreement with their corresponding tropospheric trends. In a qualitative sense, and keeping in mind the regional nature of these measurements and the uncertainties related to the calculation of stratospheric transport times, this would point towards increasing mass fluxes of CFC-11 being transported back to the troposphere. Our observations therefore do support the hypothesis of new emissions being lower than expected from tropospheric trends alone. More generally, there is evidence for a significant and time-dependent role of the stratosphere in the modulation of tropospheric trends of trace gases. However, any further quantification of the stratospheric part of the CFC-11 story is prevented firstly by the non-global and intermittent nature of sufficiently precise observations as well as their limited comparability to model or reanalysis results; secondly by the variability of the CFC-11 stratosphere-to-troposphere mass flux influenced by, for example, QBO, ENSO (El Niño–Southern Oscillation), volcanic eruptions, and also stratospheric transport changes as indicated by the observed trace gas trends on AoA surfaces; and thirdly by the large differences between results from different current meteorological reanalyses. The quality of the latter is currently the main limitation to refining such calculations.

Finally, our observations span 10 years, which is a short time in comparison to the long-term climate-change-driven stratospheric circulation changes expected from global models, which are of the order of decades (Polvani et al., 2018). Our data, however, demonstrate the capabilities of the AirCore observations to increase data coverage and better constrain such changes on various timescales.

Data availability. Observational data are included in the Supplement, and the CLaMS model data may be requested from the corresponding author.

Supplement. The supplement related to this article is available online at: https://doi.org/10.5194/acp-20-9771-2020-supplement.

Author contributions. JCL conducted the analysis of the overall dataset, participated in several campaigns, carried out some of the measurements, and led the writing of the article. ECLE, BB, HC, ESD, PH, RK, AJH, AL, SR, CS, MT, ET, and WTS contributed to the design of the AirCore and subsampling equipment and the various balloon campaigns with ECLE, ESD, and ET also involved in the halocarbon measurements and data analysis. CAMB and DEO were responsible for CARIBIC, and TR was responsible for the Geophysika aircraft measurement, sampling equipment, and related discussions. SAM provided NOAA northern hemispheric time series and useful respective insights, and JUG and FP led the modelling analysis. All authors contributed to the writing process of the article and scientific discussions surrounding it.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. This work was funded by the ERC project EXC3ITE and the UK Natural Environment Research Council. David E. Oram also received support from the National Centre for Atmospheric Science. We gratefully acknowledge the computing time for the CLaMS simulations granted on the supercomputer JURECA at Jülich Supercomputing Centre (JSC) under the VSR project ID JICG11. We thank all who helped with the balloon launches in Finland and the UK, the numerous NOAA station personnel and site scientists for sample flask collection and measurement, and Michel Bolder for collecting the Geophysika air samples; we also acknowledge the work of the Geophysika aircraft team. Related funding came from the European Space Agency (ESA, PremierEx and FRM4GHG projects), Forschungszentrum Jülich, the European Commission (FP7 projects RECONCILE, StratoClim, and H2020 project RINGO). We further thank Paul Konopka for carrying out some of the CLaMS simulations used here, Jörn Ungermann for help with code translations, and Rolf Müller for useful discussions.

Financial support. This research has been supported by the European Research Council (grant no. EXC3ITE (678904)), the Natural Environment Research Council (grant nos. NE/J021918/1 and NE/L002582/1), the Helmholtz Association (grant no. VH-NG-1128), the European Commission (grant nos. StratoClim-603557-FP7-ENV-2013-two-stage and RECONCILE-226365-FP7-ENV-2008-1), and the Dutch Science Foundation (NWO) (grant no. 865.07.001).

The article processing charges for this open-access
publications were covered by a Research Centre of the Helmholtz Association.

**Review statement.** This paper was edited by Peter Haynes and reviewed by two anonymous referees.

**References**


https://doi.org/10.5194/acp-20-9771-2020
Atmos. Chem. Phys., 20, 9771–9782, 2020


