

# Multi-model study of chemical and physical controls on transport of anthropogenic and biomass burning pollution to the Arctic

**S. A. Monks<sup>1</sup>, S. R. Arnold<sup>1</sup>, L. K. Emmons<sup>2</sup>, K. S. Law<sup>3</sup>, S. Turquety<sup>4</sup>, B. N. Duncan<sup>5</sup>, J. Flemming<sup>6</sup>, V. Huijnen<sup>7</sup>, S. Tilmes<sup>2</sup>, J. Langner<sup>8</sup>, J. Mao<sup>9</sup>, Y. Long<sup>4</sup>, J. L. Thomas<sup>3</sup>, S. D. Steenrod<sup>5</sup>, J. C. Raut<sup>3</sup>, C. Wilson<sup>1</sup>, M. P. Chipperfield<sup>1</sup>, G. S. Diskin<sup>10</sup>, A. Weinheimer<sup>2</sup>, H. Schlager<sup>11</sup>, and G. Ancellet<sup>3</sup>**

<sup>1</sup>Institute for Climate and Atmospheric Science, School of Earth and Environment, University of Leeds, Leeds, UK

<sup>2</sup>National Center for Atmospheric Research, Boulder, Colorado, USA

<sup>3</sup>Sorbonne Universités, UPMC Univ. Paris 06, Université Versailles St-Quentin, CNRS/INSU, LATMOS-IPSL, Paris, France

<sup>4</sup>Laboratoire de Météorologie Dynamique, IPSL, CNRS, UMR8539, 91128 Palaiseau Cedex, France

<sup>5</sup>NASA Goddard Space Flight Center, Greenbelt, USA

<sup>6</sup>European Centre for Medium range Weather Forecasting, Reading, UK

<sup>7</sup>Royal Netherlands Meteorological Institute (KNMI), De Bilt, the Netherlands

<sup>8</sup>Swedish Meteorological and Hydrological Institute, 60176 Norrköping, Sweden

<sup>9</sup>Program in Atmospheric and Oceanic Sciences, Princeton University and Geophysical Fluid Dynamics Laboratory/National Oceanic and Atmospheric Administration, Princeton, New Jersey, USA

<sup>10</sup>Chemistry and Dynamics Branch, NASA Langley Research Center, Hampton, VA 23681

<sup>11</sup>Deutsches Zentrum für Luft- und Raumfahrt (DLR), Institute of Atmospheric Physics, Oberpfaffenhofen, Germany

Correspondence to: S. A. Monks ([s.a.monks@leeds.ac.uk](mailto:s.a.monks@leeds.ac.uk))

## Abstract

Using observations from aircraft, surface stations and satellite, we comprehensively evaluate multi-model simulations of carbon monoxide (CO) and ozone (O<sub>3</sub>) in the Arctic and over lower latitude emission regions, as part of the POLARCAT Model Inter-comparison Project (POLMIP). Evaluation of eleven atmospheric models with chemistry shows that they generally underestimate CO throughout the Arctic troposphere, with the largest biases found during winter and spring. Negative CO biases are also found throughout the Northern Hemisphere, with multi-model mean gross errors (9–12 %) suggesting models perform similarly over Asia, North America and Europe. A multi-model annual mean tropospheric OH ( $10.8 \pm 0.6 \times 10^5$  molec. cm<sup>-3</sup>) is found to be slightly higher than previous estimates of OH constrained by methyl chloroform, suggesting negative CO biases in models may be improved through better constraints on OH. Models that have lower Arctic OH do not always show a substantial improvement in their negative CO biases, suggesting that Arctic OH is not the dominant factor controlling the Arctic CO burden in these models. In addition to these general biases, models do not capture the magnitude of CO enhancements observed in the Arctic free troposphere in summer, suggesting model errors in the simulation of plumes that are transported from anthropogenic and biomass burning sources at lower latitudes. O<sub>3</sub> in the Arctic is also generally underestimated, particularly at the surface and in the upper troposphere. Summer O<sub>3</sub> comparisons over lower latitudes show several models overestimate upper tropospheric concentrations.

Simulated CO, O<sub>3</sub> and OH all demonstrate a substantial degree of inter-model variability. Idealised CO-like tracers are used to quantitatively compare the impact of inter-model differences in transport and OH on CO in the Arctic troposphere. The tracers show that model differences in transport from Europe in winter and from Asia throughout the year are important sources of model variability at Barrow. Unlike transport, inter-model variability in OH similarly affects all regional tracers at Barrow. Comparisons of fixed lifetime and OH-loss idealised CO-like tracers throughout the Arctic troposphere show that OH differences are a much larger source of inter-model variability than transport differences. Model OH

concentrations are correlated with H<sub>2</sub>O concentrations, suggesting water vapour concentrations are linked to differences in simulated concentrations of CO and OH at high latitudes in these simulations. Despite inter-model differences in transport and OH, the relative contributions from the different source regions (North America, Europe and Asia) and different source types (anthropogenic and biomass burning) are comparable across the models. Fire emissions from the boreal regions in 2008 contribute 33, 43 and 19 % to the total Arctic CO-like tracer in spring, summer and autumn, respectively, highlighting the importance of boreal fire emissions in controlling pollutant burdens in the Arctic.

## 1 Introduction

During the 20th century, regions poleward of 60° N have warmed at a rate 50 % greater than the Northern Hemisphere (NH) average (0.09 °C decade<sup>-1</sup> compared to 0.06 °C decade<sup>-1</sup>) (ACIA, 2005). This is due to feedback mechanisms, such as the snow and sea-ice-albedo feedback, where melting ice leads to increased absorption of solar radiation, which further enhances warming in the Arctic (Serreze and Francis, 2006). Studies have shown that Arctic temperatures respond to both local and non-local radiative forcing, meaning that reducing emissions of greenhouse gases and their precursors throughout the Northern Hemisphere will be beneficial in slowing Arctic warming (Shindell, 2007). Shindell and Faluvegi (2009) estimated that anthropogenic emissions of black carbon (BC) and secondary production of tropospheric ozone (O<sub>3</sub>) from anthropogenic precursor emissions have contributed 0.5–1.4 and 0.2–0.4 °C to Arctic warming since 1890, respectively. The short lifetimes of these species compared to carbon dioxide (CO<sub>2</sub>), mean that emission reduction measures that target these species in conjunction with CO<sub>2</sub> could offer a more immediate abatement of climate warming than measures targeting solely CO<sub>2</sub> (Quinn et al., 2008). Consequently, there is a need to better understand the sources and sinks of these species and their effects on Arctic climate.

The discovery of springtime Arctic haze led to research efforts that aimed to provide insight into this phenomenon in a seemingly clean remote region of the globe. Analysis

of aircraft data from the Arctic haze and the Arctic Gas and Aerosol Sampling Program (AGASP) (Schnell and Raatz, 1984) and AGASP-II (Schnell et al., 1989), along with surface measurements showed that sulphate aerosols were the most abundant component of the haze, however, concentrations of chemically important trace gases such as carbon monoxide (CO), CO<sub>2</sub>, methane (CH<sub>4</sub>), O<sub>3</sub> and peroxyacetyl nitrate (PAN), were also elevated within the haze layers (e.g. Khalil and Rasmussen, 1984; Bottenheim et al., 1986; Bottenheim and Gallant, 1989; Solberg et al., 1996). Early source contribution studies focused on understanding the sources of sulphate aerosols due to their predominance in the haze. These studies suggested that Arctic haze was a product of inefficient removal processes in the stable cold Arctic atmosphere coupled with efficient transport of anthropogenic emissions during winter and spring from Eurasia to the Arctic (Rahn, 1985; Barrie, 1986; Barrie et al., 1989; Iversen, 1989; Shaw, 1995).

More recently, studies have begun to investigate the origins of many chemical species in addition to sulphate throughout the year (e.g. Klonecki et al., 2003; Koch and Hansen, 2005; Fisher et al., 2010; Hirdman et al., 2010; Walker et al., 2012; Wespes et al., 2012). The TOPSE (The Tropospheric Ozone Production about the Spring Equinox) aircraft campaign in 2000 provided measurements that were used to better understand Arctic O<sub>3</sub> chemistry during spring (Atlas et al., 2003). Arctic atmospheric composition in the summer and autumn has received less attention compared to winter and spring due to more effective pollutant removal mechanisms and less efficient poleward transport from the mid-latitudes (Barrie, 1986). However, it has been shown that emissions can still be regularly transported into the Arctic at this time of the year (e.g. Brock et al., 1989; Stohl, 2006; Sodemann et al., 2011). Arctic trace gases are likely to exhibit different regional sensitivities to that of aerosol due to different emission sources, longer lifetimes, and the fact that some species are not efficiently lost by wet deposition during long-range transport. This was demonstrated by Shindell et al. (2008) who examined source contributions from different NH anthropogenic emission regions to Arctic burdens of sulphate, BC, CO and O<sub>3</sub> and found that sulphate and BC were mostly sourced from Europe, in agreement with previous studies. However, for CO

and O<sub>3</sub>, North America, Asia and Europe were all shown to be important throughout the year, with North America dominating the O<sub>3</sub> burden and Europe dominating the CO burden.

The sensitivity of the Arctic to different source regions is related to their regional export efficiencies and boundary layer export mechanisms. Stohl (2006) used a Lagrangian model to study seasonal transport to the Arctic from North America, Europe and Asia and identified three main pathways that varied in terms of importance depending on the source region and season. These were low-level transport followed by ascent in the Arctic, low-level transport alone, and uplift outside the Arctic followed by descent within the Arctic. Stohl (2006) showed that European emissions could follow all three pathways, however, North American and Asian emissions tended to follow the latter. Asia and North America are ideally located to allow frequent rapid transport from the boundary layer into the free troposphere by warm conveyor belts, which occur ahead of a passing cold front (Stohl, 2001). In contrast, warm conveyor belts are not as common over Europe, and export of emissions occurs mostly at low-levels (Stohl, 2001; Duncan and Bey, 2004; Eckhardt et al., 2004). These different transport pathways result in a vertical gradient in Arctic sensitivities to different emission regions, with the higher altitudes being more sensitive to emissions from North America and Asia and the lower altitudes being more sensitive to emissions from Europe and Siberia (Klonecki et al., 2003).

Long-term changes in regional emissions may also be important in determining the relative contributions to Arctic pollutant burdens. Arctic surface observations showed a decrease in O<sub>3</sub> concentrations during the 1980s to the mid-1990s and a small increase afterwards (Oltmans et al., 1998, 2006; Helmig et al., 2007). In addition, Arctic CO surface measurements showed a downward trend during the late 1980s and early 1990s (Novelli et al., 1994; Khalil and Rasmussen, 1994). A modelling study showed a 0.5–1.1 % year<sup>-1</sup> decrease in CO between 1988–1997 in the Arctic could be explained by a decrease in European emissions (Duncan and Logan, 2008). However, total energy consumption in Asia is estimated to have more than doubled since the 1980s, resulting in a large increase in Asian emissions (Ohara et al., 2007). This is likely to have important implications for the Arctic.

There is also increasing evidence that biomass burning is an important source of Arctic pollutants during summer and autumn (e.g. Legrand et al., 1992; Wofsy et al., 1992; Harriss et al., 1994; Paris et al., 2009; Warneke et al., 2010; Bourgeois and Bey, 2011; Tilmes et al., 2011; Thomas et al., 2013), and that these emissions can influence the inter-annual variability of Arctic CO (Monks et al., 2012). Forest fires emit large quantities of O<sub>3</sub> precursors (CO, NO<sub>x</sub>, non-methane hydrocarbons (NMHC)) and aerosols, and have the potential to affect the radiative balance of the atmosphere (Randerson et al., 2006).

Understanding the relative importance of biomass burning and anthropogenic sources to Arctic burdens of short-lived climate forcers and our ability to predict future Arctic climate relies heavily on our ability to model Arctic composition and long-range transport from source regions. Shindell et al. (2008) evaluated a number of different chemical transport models (CTMs) against surface observations and highlighted that model simulated Arctic tropospheric concentrations of CO and O<sub>3</sub> were highly variable between models, and often in poor agreement with observations from Arctic surface stations. Model discrepancies were attributed to differences in model emissions, chemistry and transport. In terms of skill, a common feature of all models was a negative bias in simulated CO during winter and spring, which has also been shown to occur at the surface throughout the NH (Shindell et al., 2006). Investigation into the causes of this consistent negative bias has resulted in several possible explanations. Kopacz et al. (2010) argued that the winter/spring bias in CO could be reduced by introducing a seasonal cycle into the anthropogenic emissions, in contrast to Mao et al. (2013) who argued that a new transition metal catalysed HO<sub>2</sub> uptake channel, which acts as a loss of HO<sub>x</sub> from the atmosphere, needs to be included in models in order to increase concentrations of winter/spring CO. The generally poor skill and large variability of global models in simulating Arctic burdens of these species has implications for our confidence in a model's ability to accurately simulate climate responses to changes in mid-latitude emissions. Therefore, there is a need to better understand the chemical and physical processes that lead to these model differences and biases.

An improvement in this understanding is largely limited by a paucity of observations of chemical constituents throughout the Arctic troposphere, particularly away from the sur-

face, where enhancements from more southerly sources in North America and Asia are more likely to be found (Klonecki et al., 2003). The POLARCAT (POLAr study using Aircraft, Remote Sensing, surface measurements and models of Climate, chemistry, Aerosols, and Transport) campaign intensively sampled the Arctic and sub-Arctic during spring and summer 2008, as part of the International Polar Year (Law et al., 2014). The POLARCAT Model Intercomparison Project (POLMIP) aims to exploit this observational dataset to evaluate models in the Arctic and sub-Arctic (Emmons et al., 2014). In this study, we use POLMIP simulations of CO, O<sub>3</sub> and OH in conjunction with POLARCAT aircraft observations, surface measurements and satellite observations, within the Arctic and near source regions to evaluate model performance (Sect. 4). The availability of the aircraft observations in the Arctic allows a more extensive multi-model evaluation of models within a variety of air-mass types in the Arctic than previously undertaken. Idealised tracers are used to provide a useful summary of the relative importance of anthropogenic and biomass burning sources from several models throughout the year 2008 and to investigate inter-model differences in these source contributions (Sect. 5). The aim of this paper is to provide a better understanding of inter-model differences in simulated Arctic trace gases.

## 2 POLMIP model simulations

A total of nine global and two regional (WRF-Chem and MATCH) three-dimensional models with chemistry were run for 2008 as part of POLMIP. Table 1 shows the participating models' horizontal/vertical resolutions, number of gas-phase species and reactions and meteorology used. More details on the individual models and the set-up of the POLMIP experiments are given in Emmons et al. (2014). Most models were run offline, forced by either ECMWF or GEOS-5 meteorological data (see Table 1). The meteorological fields of C-IFS are re-initialised with ECMWF-Operational analyses every 24 h, calculating the meteorology online in between. LMDZ-INCA uses the offline ECMWF winds but calculates other meteorological data online (e.g. temperature and humidity). The WRF-Chem model was run at two different horizontal resolutions (100 and 50 km) to show the impact of resolution on Arctic



simulations of chemistry and transport. Due to the computational expense of this model, these simulations were performed over a region and time period covering only the summer POLARCAT flights (using the MOZART concentration fields as boundary conditions).

Anthropogenic, natural and biomass burning emissions were specified for the POLMIP experiments on a  $1^\circ \times 1^\circ$  grid. Monthly mean anthropogenic and ship emissions were based on the Streets v1.2 inventory (Zhang et al., 2009), which was updated with the latest regional inventories in 2008 for the POLARCAT campaign. Monthly mean natural emissions from the MACC (Monitoring Atmospheric Composition and Climate) project (MACC-ity), which are based on simulations from the Model of Emissions of Gases and Aerosols from Nature (MEGAN) v2.1 (Guenther et al., 2012) were used. For the comparison period of 2008, daily biomass burning emissions were used from the Fire INventory from NCAR (FINN) (Wiedinmyer et al., 2011). These were emitted into the surface level of each model, with the exception of WRF-Chem, which used a fire plume model. It should be noted that the GEOS-Chem model did not use the same anthropogenic emission datasets as described here, with emissions of CO being approximately 10% lower than in other models and propane being a factor of two higher (Emmons et al., 2014). For this reason, it should be remembered that this model is not directly comparable. Monthly mean output for 2008 and hourly instantaneous output for the duration of the POLARCAT spring and summer campaign periods are used in this study. Models are excluded from comparisons if the required data is not available.

Some models also included six regional tracers with prescribed constant atmospheric lifetimes of 25-days, to allow model differences in transport to be examined. These tracers were emitted from the regions of North America, Europe and Asia, with a tracer for both fire and anthropogenic-sourced CO in each of the 3 regions (see Emmons et al. (2014) for details). For this study, a further set of these six tracers were calculated in the TOMCAT model, using the same emissions, with loss determined by monthly mean OH concentration fields taken from each POLMIP model, and kinetics of CO + OH loss ( $k = 1.44 \times 10^{-13}(1 + [N_2])/4.2 \times 10^{19}$ ). These tracers therefore had the same transport, as calculated by TOMCAT, but different loss rates due to the different OH fields, allowing the impact of OH differences

on Arctic burdens of CO in the models to be examined through their differences in OH. CO and CH<sub>4</sub> tracers with loss determined by offline OH have been used within the TOMCAT model previously (Monks et al., 2012; Patra et al., 2011). The models that took part in the tracer experiments are shown in Table 1.

### 3 Observations

A range of observations are used to evaluate the POLMIP model simulations on different temporal and spatial scales. The measurement uncertainties and techniques have been summarised in Table 2. Monthly mean observed surface CO and O<sub>3</sub> at stations in the Arctic are taken from the World Data Centre for Greenhouse Gases (WDCGG) (<http://ds.data.jma.go.jp/gmd/wdcgg/>) provided by NOAA/ESRL. CO is from flask samples analysed using gas chromatography (Novelli et al., 1998) and O<sub>3</sub> is measured by ultra-violet (UV) light absorption at 254 nm (Oltmans and Levy, 1994). For model-observation comparisons, the model simulated CO and O<sub>3</sub> concentrations are interpolated both horizontally and vertically to the location of the observatory (see Fig. 1 for station locations).

Simulated CO is compared on a global and regional scale to CO retrieved from the satellite instrument, MOPITT (Measurements Of Pollution In The Troposphere) version 6. MOPITT is a nadir-viewing instrument on-board the NASA Terra satellite and retrieves global concentrations of CO at a horizontal resolution of  $\sim 22$  km by measuring infrared radiances in the CO absorption band (Deeter et al., 2010). Version 6 is the latest release and uses an updated a priori based on a climatology of 2000–2009 output from the CAM-Chem model (Deeter et al., 2014). It also has an increased sensitivity to lower tropospheric CO by exploiting measurements in the near-infrared and thermal infrared (Deeter et al., 2011). As with all nadir-viewing instruments, MOPITT is more sensitive to certain altitudes, therefore averaging kernels that contain information about the instrument's varying sensitivities at different altitudes are supplied with the retrieved data. These are used, along with the a priori, to apply the vertical sensitivity of the satellite instrument to the simulated CO profiles from the POLMIP models. This allows a more accurate comparison between the observed and

simulated CO. Version 6 of the MOPITT retrieval has been shown to have a bias of between -5% to 9%, with the largest positive bias occurring at the surface and the largest negative bias occurring at 400 hPa (Deeter et al., 2014).

MOZAIC (Measurement of O<sub>3</sub> and water vapour by Airbus In-service airCRAFT) data is collected on-board Airbus commercial aircraft during ascent from and descent to airports, offering detailed profiles of CO and O<sub>3</sub> near source regions. The June/July model hourly output is interpolated both horizontally and vertically to the MOZAIC aircraft locations in order to examine the vertical structure of these gases over Europe and North America. Unfortunately, no MOZAIC data is available within any of the required comparison regions in April 2008 or over Asia in June/July 2008. The number and location of MOZAIC profiles made during June–July 2008 over Europe and North America are shown in Fig. 1.

The POLARCAT project brought together several international groups of scientists and coordinated intensive measurement campaigns in different regions of the Arctic and sub-Arctic during April and June–July 2008 (Law et al., 2014). These included the ARCTAS (Arctic Research of the Composition of the Troposphere from Aircraft and Satellites), POLARCAT-GRACE and POLARCAT-France aircraft campaigns. The DC8 aircraft was based in Fairbanks, Alaska from 1st–21 April 2008 during ARCTAS-A and then in Cold Lake, Canada from 29 June–10 July 2008 during ARCTAS-B (Jacob et al., 2010). The POLARCAT-GRACE and POLARCAT-France projects had two aircraft jointly based in Kangerlussuaq, Greenland in summer. Even though based at the same location, the Falcon had a larger range than the ATR-42 and was able to cover larger regions and higher altitudes. The POLARCAT flight tracks are shown in Fig. 1.

#### 4 Model evaluation

In this section, POLMIP model simulations are evaluated against Arctic surface CO and O<sub>3</sub> data and MOPITT CO retrievals throughout the year 2008. Aircraft measurements of CO, O<sub>3</sub> and OH from the POLARCAT and MOZAIC projects are then used to evaluate the vertical structure of the troposphere during intensive periods of sampling in April and June–

July 2008. Model performance against the observations is summarised in Fig. 10 using the normalised mean gross error (NMGE). This gives the mean model bias (regardless of sign) over the vertical profile (when comparing to aircraft data) or over the whole year at a certain altitude (when comparing against surface and MOPITT data) as a percentage of the observed concentrations.

#### 4.1 Arctic surface comparisons of carbon monoxide and ozone

Fig. 2 shows the time series of monthly mean 2008 simulated and observed concentrations of CO at Barrow and Zeppelin and O<sub>3</sub> at Barrow and Summit (stations located north of the Arctic Circle). These models use the same emissions data, removing one key inter-model difference in Arctic surface CO and O<sub>3</sub> comparisons identified in Shindell et al. (2008). The overall model performance at several Arctic stations (shown in Fig. 1) is summarised as Taylor diagrams in Fig. 3. Perfect agreement between a model and observations would result in a normalised standard deviation (NSD) of 1, a correlation ( $r$ ) of 1 and a centered root-mean-square difference (RMSD) of 0 (indicated by “Observed” on the Taylor diagram).

The models capture the seasonality in CO, with correlations greater than 0.8 at all surface stations. The models show a large range in their ability to capture the amplitude of the observed seasonal cycle ( $0.3 \leq \text{NSD} \leq 1.2$ ). RMSDs are mostly clustered between values of 1 and 2. As this error statistic is also a function of the monthly deviations from the annual mean (see Fig. 3 caption), the models with a higher RMSD are the models that do not capture the amplitude of the seasonal cycle well. Fig. 4 shows box and whisker plots of simulated seasonal mean surface CO bias at Barrow and Zeppelin. The median biases tend to lie below the zero line, showing that models generally underestimate CO throughout the year. However, the median biases are near zero in autumn, with some models overestimating CO in summer and autumn. The largest negative median biases occur during winter at Barrow and during spring at Zeppelin, with the smallest median biases occurring during autumn at both stations. This shows that state-of-the-art models still consistently underestimate winter/spring Arctic surface CO as shown in previous studies (e.g. Shindell et al., 2006, 2008).

Two models, GEOS-Chem and LMDZ-INCA, stand out from the other models in winter/spring, showing much better agreement with the observations due to higher simulated CO concentrations compared to other models. The version of the GEOS-Chem model used in this study includes a new scheme for transition metal catalysed heterogeneous loss of HO<sub>2</sub> on aerosol that produces water rather than H<sub>2</sub>O<sub>2</sub>, which is particularly important in spring in the Northern Hemisphere (Mao et al., 2013). The scheme results in lower HO<sub>x</sub> concentrations, which increases the atmospheric lifetime of CO and may explain the improved simulation of winter/spring Arctic surface CO by GEOS-Chem. The LMDZ-INCA model has lower OH compared to the other models (see Sect. 5.1), likely explaining the higher simulated CO concentrations. The C-IFS and MATCH models, which are run at the two highest horizontal resolutions (see Table 1), show no improvement over the other models in the simulation of CO.

The Shindell et al. (2008) model intercomparison study calculated root-mean-square errors (RMSE) averaged over two Arctic surface stations (Barrow and Alert) of 17–40 ppbv for CO, with two models having particularly high RMSE of 54 and 83 ppbv. For the POLMIP models, RMSE values of 13–33 ppbv are calculated for the same two stations. Shindell et al. (2008) showed a larger spread in model summertime surface CO compared to that shown in Fig. 2. Direct comparison between the POLMIP models and Shindell et al. (2008) is difficult due to different participating models and changes in model set-up and emissions. However, it may suggest that inter-model differences are reduced in summer due to the use of a single emissions dataset in most of the POLMIP models.

For O<sub>3</sub>, the Taylor diagram shows that there is a much larger spread in model performance than for CO ( $0.1 \leq r \leq 0.9$ ,  $0.5 \leq \text{NSD} \leq 2.5$  and  $1 \leq \text{RMSD} \leq 9$ ). It is also clear that all models perform markedly worse at Barrow compared to Summit. At Barrow, observations show a decrease in O<sub>3</sub> in March due to well-known halogen-induced ozone depletion events (ODE) in the boundary layer at this location (Barrie et al., 1988; Helmig et al., 2007). In contrast, most models simulate increasing concentrations between February and April, most likely due to the lack of halogen chemistry in the models, resulting in the low correlations at Barrow. The GEOS-Chem model, which does include halogen chemistry, is still

not able to fully capture the seasonal transition between February and April even though in March the absolute concentrations agree well. This model also simulates  $O_3$  that is lower than the other models and observations between late-spring/early-summer, resulting in a larger negative bias against observations compared to other models. It is not clear what is causing this bias in the model.

Median biases of  $O_3$  at these two stations (see Fig. 4) show that the models generally underestimate  $O_3$  throughout the year at the surface, with the exception of spring at Barrow, in agreement with Shindell et al. (2008). The POLMIP models have RMSE of 5–10 ppbv averaged over Summit and Barrow, which is slightly lower than the RMSE of 7–12 ppbv from Shindell et al. (2008). At Barrow, the largest positive median bias occurs during spring (due to ODE) and the largest negative median bias occurs during winter. In contrast, the largest negative bias at Summit occurs during spring, closely followed by summer. It should also be noted that, whilst simulated  $O_3$  seasonality at Summit shows much better agreement with the observations, the absolute median model bias is larger compared to that at Barrow. The high altitude of Summit (3238 m) makes this station highly sensitive to stratospheric ozone enriched air (Helmig et al., 2007; Hirdman et al., 2010), explaining the higher  $O_3$  concentrations. Therefore the higher absolute bias at this station may indicate an underestimate in the stratospheric source of  $O_3$ .

## 4.2 MOPITT carbon monoxide comparisons

Fig. 5 shows the monthly average MOPITT total CO column for April and July 2008 along with the difference between the multi-model mean and MOPITT total column. The multi-model mean was calculated from ten POLMIP model total columns with the averaging kernels applied. Missing data and retrievals where the degrees of freedom signal (DOFS) are less than 1 have been removed from the model and satellite columns.

The highest concentrations in April are observed over eastern Asia near China. This high CO is a result of Asian anthropogenic emissions combined with unusually early and large fires in southern Siberia during April 2008 (Warneke et al., 2010; Brock et al., 2011; Law et al., 2014). MOPITT also exhibits high CO over parts of Europe and North

America due to fossil fuel emissions and over western Russia due to agricultural burning (Warneke et al., 2009). It is also possible to see the large-scale continental outflow over the Pacific and Atlantic. The CO concentrations seen by MOPITT are lower in July due to a shorter atmospheric lifetime, however, regions of high CO can still be seen over source regions and the oceans due to long-range transport. In particular, wildfires that occurred during the summer of 2008 are visible in eastern Siberia and Canada. Similar fire signatures were also observed by the Infrared Atmospheric Sounding Interferometer (IASI) satellite instrument during spring and summer (Pommier et al., 2010), however MOPITT total column background concentrations are slightly lower in general. Along with CO, plumes originating from these sources were elevated in other important trace gases and aerosols, and on several occasions they were transported to the Arctic embedded within frontal systems (Fuelberg et al., 2010). These were then subsequently sampled by aircraft as part of POLARCAT campaign (Law et al., 2014).

The multi-model mean total column CO shows a widespread negative bias throughout the Northern Hemisphere in both April and July. The bias is clearly larger in spring compared to summer, which is in agreement with the Arctic surface comparisons (in Sect. 4.1) and previous studies (e.g. Shindell et al., 2006). A large positive model bias is seen over south-east Asia in April. Shindell et al. (2006) compared multi-model simulations to MOPITT in April 2000–2004 and did not show a similar positive bias in this region. This suggests that the bias seen in the POLMIP models in 2008 may be specific to the year considered. As already mentioned, there was an unusually early start to the fire season in April 2008 with extensive burning occurring in eastern Siberia. The NCAR FINN fire emission inventory shows high CO emissions over Siberia and Myanmar (Wiedinmyer et al., 2011). The emissions over Myanmar are co-located with the positive model bias, which suggests the fire emissions are overestimated in this region. MOPITT observes elevated CO over Myanmar, however, a more extensive region of high CO is seen further North over north-east China and south-east Russia. The overestimation of emissions over Myanmar results in a plume of elevated CO being simulated over the Pacific at more southerly latitudes than observed by MOPITT.

Fig. 6 shows monthly mean MOPITT CO at 700 and 300 hPa averaged over the Arctic and the three major NH emission regions (North America, Europe and Asia), allowing model performance over different source regions to be examined. The mean DOFS over the Arctic region (see top panel of Fig. 6) is not much more than 1, indicating the MOPITT retrievals have little vertical information in the Arctic, which is due to the lack of thermal contrast between the ground and the atmosphere. This makes comparison between model performance in the Arctic and over the source regions difficult. At 700 hPa, all models capture the seasonal cycle reasonably well, both in the Arctic and over the source regions ( $r = 0.94$ – $0.99$ ). At 300 hPa, a larger range in correlations ( $r = 0.47$ – $0.98$ ) indicates that the models are less able to capture the seasonal cycle at this altitude. As seen at the Arctic surface, most models underestimate CO at both levels throughout the year. The models which have the lowest/highest concentrations of CO throughout the year in the Arctic also have the lowest/highest concentrations near the source regions suggesting the individual model biases are consistent throughout the NH. The monthly model biases show a seasonal cycle (see bottom panel, Fig. 6), which is present across the different regions and models. At 700 hPa the largest regional biases tend to occur in winter and spring, as found at the Arctic surface stations, and the smallest biases occur in summer. However, at 300 hPa the bias shows the opposite behaviour, where the largest bias in all regions occurs in summer. This suggests that the surface winter/spring bias may be shifted to higher altitudes during summer when vertical mixing of pollutants by convection is important (e.g. Hov and Flatøy, 1997; Donnell et al., 2001; Jaegle, 2007).

The multi-model normalised mean gross error (NMGE, shown in Fig. 10), is found to show different vertical sensitivities depending on the region. Over North America, the NMGE is smaller at 700 hPa relative to 300 hPa, whilst over Europe the NMGE is lower at 300 hPa relative to 700 hPa. In contrast, the NMGE is similar at both 300 hPa and 700 hPa over Asia. However, the overall range in error is relatively small (9–12%).



### 4.3 POLARCAT Arctic aircraft comparisons

Aircraft measurements from the spring and summer 2008 POLARCAT campaign allow a detailed insight into model performance in the Arctic throughout the troposphere over regions of Greenland and the Canadian Arctic. The hourly gas-phase species from each model have been linearly interpolated in time and space to the location of the aircraft. The observed and simulated concentrations are separated into 50 hPa bins to give average vertical profiles over all flights. Aircraft profiles of CO, O<sub>3</sub>, OH and water vapour for the ARCTAS-A (spring) and ARCTAS-B (summer) campaigns are shown in Fig. 7 and profiles of CO and O<sub>3</sub> for the POLARCAT-France and POLARCAT-GRACE summer campaigns are shown in Fig. 8.

#### 4.3.1 ARCTAS-A spring comparisons

In April 2008, the CO and O<sub>3</sub> DC8 observations show a fairly homogeneous profile, with only small changes in concentrations from the surface up to the middle troposphere. Around 50 % of the CO measured came from anthropogenic sources in Asia, North America and Europe, with Asian emissions dominating (25 %) (Bian et al., 2013), most of which was emitted in China and India (Tilmes et al., 2011). Biomass burning emissions also contributed to the sampled CO, causing small enhancements in the observed profile (Bian et al., 2013; Tilmes et al., 2011) that can be seen at around 500 and 750 hPa in Fig. 7. Biomass burning also contributed a few ppbv to the sampled O<sub>3</sub> at the same levels (Wespes et al., 2012). In the upper troposphere, the influence of stratospheric mixing is at its greatest (Wespes et al., 2012), which can be seen by a large increase in O<sub>3</sub> and a concurrent decrease in CO.

The models capture the vertical structure well, yielding correlations of 0.96–0.99 for CO and 0.88–1.00 for O<sub>3</sub>, however, only very small CO enhancements are seen in the simulated profiles at the altitude of the boreal fire plumes. This may be due to simulated fire emissions being exported from Asia at more southerly latitudes than observed by MOPITT (see Sect. 4.2). This will affect the transport of emissions to Alaska and Canada, where the DC8 flights took place and subsequently sampled these plumes. Most models show nega-

tive biases in CO from the surface up to around 300 hPa, which are similar in magnitude to those seen in Fig. 2 in April. LMDZ-INCA and GEOS-Chem stand out from the other models as seen at the Arctic surface (Fig. 2, Sect. 4.1), showing higher simulated CO up into the free troposphere. This results in better agreement with the aircraft measurements. In the upper troposphere, most models show good agreement with the observations.

As with CO, most models underestimate O<sub>3</sub> concentrations in the lower troposphere and then show good agreement in the upper troposphere.

The observed OH shows a much more vertically heterogeneous profile due to its very short lifetime compared to CO and O<sub>3</sub>, with maximum concentrations being found in the upper troposphere. The ability of the models to capture this vertical structure is highly varied, with correlations ranging from 0.63 to 0.98. However, OH measurements have a particularly high uncertainty (see Table 2) and the short lifetime of OH makes comparisons with coarse global models difficult. The simulated OH concentrations vary greatly across the models, with some models having almost twice as much OH throughout the troposphere compared to other models. The models with lower OH tend to simulate OH concentrations that are slightly lower than observed throughout the troposphere. The models with the higher OH tend to simulate concentrations which are slightly higher than observed in the upper troposphere. In the lower and middle troposphere, these models simulate median OH concentrations lower than the observed median concentrations but lie within the 25th and 75th percentiles. Emmons et al. (2014) showed that the models with higher OH also had higher photolysis rates of O<sub>3</sub> to O(<sup>1</sup>D) and that a subset of models (CAM4-Chem, CAM5-Chem, MOZART-4) exhibited a relationship between higher photolysis rates and lower cloud cover fraction. These are the models that show higher OH concentrations in the upper troposphere against the ARCTAS data. Inter-model differences in OH and the impact on the Arctic are discussed in more detail in Sect. 5.1.

Model water vapour concentrations show good agreement with the observations apart from in the upper troposphere where there is evidence that the models overestimate observed concentrations. This may be contributing to a possible overestimation of OH in this region in some models.

### 4.3.2 ARCTAS-B summer comparisons

In June–July 2008, the DC8 aircraft covered large regions of the Canadian Arctic, sampling a range of polluted air masses (as indicated by the large range of sampled concentrations). The aircraft was ideally located to sample local fires that were burning in Canada at the time of the campaign, resulting in  $\geq 60\%$  of the observed CO in the lower troposphere being directly emitted from boreal biomass burning (Bian et al., 2013). This can be seen in the observed profile in Fig. 7 as a large enhancement in CO at 900 hPa. Another enhancement in the observed CO profile is seen at 300 hPa, which is primarily due to the long-range transport of Asian anthropogenic and Siberian biomass burning emissions to Canada (Bian et al., 2013). O<sub>3</sub> shows a concurrent decrease in concentrations at 300 hPa and low O<sub>3</sub> near the surface, indicating O<sub>3</sub> production may have been limited within these plumes. In the lower troposphere, where fresh fire plumes were sampled, this is most likely due to the rapid conversion of NO<sub>x</sub> to PAN, limiting O<sub>3</sub> production in the fire plumes (Alvarado et al., 2010). At 300 hPa, lower O<sub>3</sub> concentrations coincided with a limited influence from stratospheric O<sub>3</sub> relative to the surrounding air, suggesting the sampling of air-masses that are not well-mixed with background air (Wespes et al., 2012).

Lower model-observed CO correlations ( $r = 0.71$ – $0.95$ ) show that the models agree less well with the observed CO vertical structure in the summer campaign compared to the spring, however, the main features of the profile are captured, with all models showing enhancements in CO at 300 and 900 hPa. Even though these enhancements are simulated by all models, the absolute concentrations vary widely, with some models having a negative bias and some models having a positive bias. In the middle troposphere, the models underestimate CO, as seen in spring, which suggests model background CO is still too low in summer in the middle troposphere. Global and regional models show similar overall performance, however the WRF-Chem 100 km resolution simulation has higher CO concentrations within the plume of enhanced CO at 900 hPa compared to the 50 km simulation, suggesting model resolution has important impacts on concentrations near emission sources. Whilst it is expected that a model run at a higher horizontal resolution would simulate higher

concentrations, due to the emissions being subjected to less numerical dilution due to the smaller spatial resolution of a 50 km grid box, the model will also become more sensitive to emission location errors where emissions may be emitted into neighbouring grid boxes. This will be particularly important for the model performance in the Arctic, where long-range transport of plumes plays an important role.

For  $O_3$ , the models show good agreement with the observed profile ( $r = 0.78$ – $0.99$ ). The models capture the decrease in  $O_3$  seen at 300 hPa and the lower concentrations near the surface, however, as with CO, the models show a wide range in simulated  $O_3$  concentrations at 300 hPa. This may be related to the different chemical mechanisms in the models, which result in different  $O_3$  production and loss efficiencies (Arnold et al., 2014). Summer OH correlations of 0.25–0.71 are lower than those calculated for the spring profile. In summer more rapid production and loss of OH leads to the observations being more sensitive to local changes in cloud cover, water vapour and  $O_3$ , making it even more difficult to simulate small spatial variability. Models show a wide spread in concentrations, however, they mostly lie within the 25th and 75th percentile of observed OH concentration. Model  $H_2O$  concentrations show good agreement with observations with the exception of a positive bias in the upper troposphere, as seen in spring. Model  $H_2O$  percent errors are slightly higher for most models in summer.

### 4.3.3 POLARCAT-France and POLARCAT-GRACE summer comparisons

As part of the summer POLARCAT campaign, the POLARCAT-France and POLARCAT-GRACE projects had two aircraft, the French ATR-42 and German DLR Falcon, based in Greenland during June–July 2008 to sample aged pollution transported to the Arctic (Law et al., 2014). For this reason, CO profiles sampled by these two aircraft (see Fig. 8) show lower concentrations compared to the DC8 summer measurements. They sampled a mixture of air types including background air, stratospheric air and plumes from anthropogenic and biomass burning sources that had been transported from North America, Asia and Europe (Law et al., 2014; Tilmes et al., 2011). Evidence of these polluted plumes can be seen in the ATR-42 CO profile at around 500 hPa and in the Falcon CO profile at 400 hPa.

As seen during the ARCTAS-B campaign, there is a decrease in Falcon observed  $O_3$  in these plumes relative to the surrounding air. Very little or no local emissions are indicated by CO concentrations that are lower than 100 ppbv being measured in the boundary layer.

As observed, modelled CO over Greenland is lower in comparison to concentrations simulated over the Canadian Arctic during the summer ARCTAS-B campaign. Most models underestimate the summer observed CO and show a large spread in simulated concentrations and their ability to capture the vertical structure ( $r = 0.19$ – $0.91$  for the ATR-42 and  $r = -0.18$ – $0.97$  for the Falcon). In addition, they do not capture the magnitude of the CO enhancements in the upper troposphere. These enhancements are largely due to transport of fire and anthropogenic emissions from North America (Thomas et al., 2013; Tilmes et al., 2011) and cross-polar transport of Asian anthropogenic and Siberian biomass burning emissions (Sodemann et al., 2011; Tilmes et al., 2011). Therefore, the underestimated CO enhancement in most models indicates that there may be an error in the transport of these plumes to the Arctic or in the emissions. Sodemann et al. (2011) showed that whilst a global Eulerian chemical transport model (TOMCAT) accurately simulates the cross-polar transport of these Asian plumes, it overly mixes the CO plumes with surrounding air compared to a Lagrangian model, resulting in lower than observed concentrations. This may explain why most models do not capture the magnitude of these CO enhancements. The high resolution regional model, WRF-Chem, does not show any improvement in the simulation of this enhanced CO plume compared to the global models. Increasing horizontal resolution within WRF-Chem (from 100 to 50 km) also has very little impact. Thomas et al. (2013) found the location of the North American plume in the regional WRF-Chem model to be located just outside of the sampling region of the ATR-42 and Falcon on certain days, indicating possible transport errors in the models that may explain the low CO enhancement.

For  $O_3$ , the range in correlations is slightly larger than those seen for the summer ARCTAS-B aircraft comparisons ( $r = 0.86$ – $0.98$  for the ATR-42 and  $r = 0.75$ – $0.99$  for the Falcon). There is some evidence that the models underestimate surface  $O_3$ . Helmig et al. (2009) showed that surface  $O_3$  concentrations in Greenland were dominated by  $O_3$  loss, not production, suggesting this negative model bias is not likely to be caused by a lack

of O<sub>3</sub> production in the models. Processes that may be contributing to model biases in this region could be overly efficient O<sub>3</sub> dry deposition on snow/ice land cover types or difficulty simulating the shallow stable Arctic boundary layer. The Falcon sampled a large amount of stratospheric air above 350 hPa, and on occasion flew within the stratosphere (e.g. Roiger et al., 2011), as shown by high concentrations of O<sub>3</sub> and low concentrations of CO. Most models underestimate O<sub>3</sub> concentrations in the upper troposphere suggesting the influence of stratospheric O<sub>3</sub> may be too low in the models. Fig. 10 shows that the multi-model O<sub>3</sub> NMGE is lower for comparisons against the ATR-42 data than it is for Falcon data. This is due to the ATR-42 sampling air that was less influenced by stratospheric air due to lower flight levels. The higher multi-model mean NMGE against the Falcon data is likely a result of the large negative bias seen in the upper troposphere, most notably in the TOMCAT, C-IFS and MATCH models. These models have a higher number of vertical levels compared to other models at this altitude suggesting vertical resolution is not the cause of the larger bias in these models.

#### 4.4 MOZAIC summer aircraft comparisons near source regions

Profiles of CO and O<sub>3</sub> sampled by MOZAIC aircraft during ascent from and descent to airports over North America and Europe between 18 June and 14 July 2008 (the duration of the POLARCAT campaign) are shown in Fig. 9. Overall, the models capture the vertical profiles well, with correlations greater than 0.61 for CO and 0.82 for O<sub>3</sub>.

For CO, models exhibit a negative bias over both of the selected regions between 2 and 6 km (~ 800–500 hPa). This bias can also be seen in most models in the MOPITT comparison at 700 hPa (see bottom panel, Fig. 6). As the models show good agreement at the surface, this negative CO bias may be related to a problem with the simulated export of primary emissions from the boundary layer to the free troposphere suggesting that model transport processes, such as convection, need to be evaluated in more detail near source regions. This bias could also be a result of missing emissions in the inventories near the flight locations.

There is no strong evidence of a larger summertime bias occurring at 9 km ( $\sim 300$  hPa) compared to 3 km ( $\sim 700$  hPa) as suggested by the MOPITT comparisons. The MOZAIC data shown here was collected at a limited number of locations, therefore the horizontal coverage is not as extensive as the MOPITT data and may explain some of the differences seen between these two observational datasets.

For  $O_3$ , the models generally lie within the 25th and 75th percentile of observed  $O_3$ , with some evidence of  $O_3$  being overestimated in the upper troposphere over Europe. Overall, model biases (NMGE) are larger over Europe compared to North America for CO and  $O_3$  (see Fig. 10).

#### 4.5 Overall model performance

Models tend to underestimate Arctic CO at the surface and throughout the free troposphere, most notably in winter/spring. Similar seasonal biases are also found at lower latitudes. MOPITT retrievals show that the smallest multi-model normalised mean gross error (NMGE) occurs over North America (9%) and the largest over Europe (12%) in the lower troposphere, whilst in the upper troposphere the smallest NMGE occurs over Europe (9%) and the largest over North America (12%). NMGE over Asia are similar in the lower and upper troposphere (9% and (10%, respectively). This suggests that models do not perform substantially worse over any one Arctic source region. However, detailed vertical profiles from the MOZAIC dataset do show that models exhibit slightly higher model biases over Europe (15%) than those over North America (13%). In Fig. 10, the multi-model NMGE for the different POLARCAT flights show that model error in the Arctic is highly dependent on the dataset used (NMGE = 9–22%). Comparisons against the POLARCAT ATR-42 flights show that the models have a similar NMGE to those calculated for the datasets used at lower latitudes. The other three POLARCAT datasets, however, show higher NMGE than any other CO comparisons performed in this study. The poor performance of models against the ARCTAS-A spring data is due to ubiquitous Arctic springtime negative CO model biases and for ARCTAS-B and POLARCAT Falcon summer comparisons, poor model performance

is largely due to models simulating smaller than observed CO enhancements within polluted air masses sampled by the aircraft.

Overall, models are found to underestimate  $O_3$  in the Arctic (NMGE = 12–19 %), particularly at the surface and in the upper troposphere. In contrast, several models overestimate  $O_3$  in summer over North America (14 %) and over Europe (24 %). These comparisons suggest that model improvements to Arctic boundary layer processes and stratospheric-tropospheric exchange of  $O_3$  may improve simulations of Arctic  $O_3$ .

In addition to the model biases, a substantial amount of inter-model variability is seen in CO,  $O_3$  and OH throughout the troposphere.

## 5 Drivers of Arctic model variability and impacts on source contributions

In this section, the relative contributions of transport and chemistry to inter-model differences in Arctic CO burdens are investigated. This is done using the regional CO-like tracers described in Sect. 2, which either have a 25-day fixed-loss rate and varying transport (provided by each model) or a varying loss rate dependent on each model's monthly mean OH concentrations and the same transport (run as tracers within the TOMCAT CTM). It should be noted that whilst this will quantify the effect of transport and chemistry on inter-model differences in the idealised CO-like tracers, the results will not directly equate to the same inter-model differences in trace gases such as CO and  $O_3$  due to other important factors, such as secondary CO production from other gases, such as  $CH_4$  and NMHC.

### 5.1 Model differences in OH

Fig. 11 shows the annual zonal mean OH concentrations from each model, which are used to calculate loss rates for the OH-loss tracers. The primary route for OH production is by photolysis of  $O_3$  in the presence of water vapour (Levy, 1971), with secondary production by recycling of  $HO_2$  and  $H_2O_2$ , where the concentrations of  $NO_x$  and CO are important factors (Logan et al., 1981). The models simulate the highest concentrations of OH in the tropics, where the amount of incoming sunlight is at its greatest and  $O_3$  is readily photolysed. Even



though the models agree on this zonal distribution, the magnitude of OH and the location of the annual mean peak in OH vary. Due to the importance of OH in controlling the lifetimes of most reactive trace gases in the atmosphere, these inter-model OH differences have important consequences for CO, both as its primary loss route and as a driver of secondary production from hydrocarbon oxidation.

Previous studies have related inter-model and intra-model changes in OH to variables which control its abundance. Voulgarakis et al. (2013) showed that present-day global air-mass weighted OH concentrations from the Atmospheric Chemistry and Climate Model Intercomparison Project (ACCMIP) models were linearly related to model photolysis rates of  $O_3$  to  $O(^1D)$  ( $JO(^1D)$ ) and total volatile organic carbon (VOC) emissions. Murray et al. (2014) found that differences in the  $S_N/(S_C^{3/2})$  ratio, where  $S_N$  and  $S_C$  are the total sources of  $NO_x$  and of CO and hydrocarbons, respectively, together with  $O_3$  photolysis rates and specific humidity explained most intra-model OH variability on glacial-interglacial time-scales.

To investigate POLMIP inter-model OH differences in the Arctic and on a global-scale, model spread in mean tropospheric air-mass weighted OH has been related to the model spread in several variables that are likely to impact OH, such as, the ratio of the total tropospheric burden of CO to  $NO_x$ ,  $J(O^1D)$ , water vapour concentrations and  $O_3$  concentrations. The coefficient of determination ( $r^2$ ) and the linear regression coefficients ( $b$ ), calculated using ordinary-least squares (OLS) and iteratively re-weighted least squares (IRLS) methods, for the full set of parameters considered are shown in Supplement Table S1.

The model variables that have the highest  $r^2$  values are shown in Fig. 12. In the Arctic, a  $r^2$  value of 0.85 is found between OH and water vapour concentrations, which is significant above the 90 % confidence level (CL). On a global scale, a  $r^2$  value of 0.35 is calculated, which is not significant at the 90 % CL, however, the OLS and IRLS regression analysis highlights MOZART-4 water vapour as an outlier. When MOZART-4 is removed from this part of the analysis, a  $r^2$  value of 0.96 is calculated (see Supplement Table S1). These results suggest that inter-model differences in  $H_2O$  concentrations in the Arctic and throughout the troposphere are an important driver of inter-model OH differences in these simulations. It is interesting to note that the models that use ECMWF reanalyses also have

some of the lowest global mean OH concentrations (C-IFS, TM5 and TOMCAT), however, C-IFS and TM5 both use the same chemical mechanism, which has been shown to produce lower OH concentrations compared to other mechanisms (Archibald et al., 2010). Further investigation into how sensitive model tropospheric OH and other trace gases, such as O<sub>3</sub> and CO, are to model differences in meteorology would be worthwhile.

In the Arctic, the  $J(\text{O}^1\text{D})$  rates multiplied by O<sub>3</sub> concentrations are found to be the second most important variable in controlling inter-model OH differences ( $r^2 = 0.65$ ), where the models with higher photolysis rates also have higher OH concentrations, as expected. On the global scale this relationship is much weaker and is not significant, in contrast to the ACCMIP chemistry-climate models (Voulgarakis et al., 2013). The global CO burden is found to be the variable with the second highest  $r^2$  value against the global tropospheric OH concentration, however, as OH controls the lifetime of CO it is not surprising to see such a correlation and it is unlikely to explain the spread in OH as the models use the same emission inventories.

Prinn et al. (2001) and Bousquet et al. (2005) estimated annual mean tropospheric air-mass weighted OH concentrations of  $9.4 \pm 1.3 \times 10^5 \text{ molec cm}^{-3}$  and  $9.8 \pm 1.3 \times 10^5 \text{ molec cm}^{-3}$ , respectively, from the lifetime of the methyl chloroform. Eight POLMIP models simulate air-mass weighted global mean OH concentrations of  $10.1\text{--}12.0 \times 10^5 \text{ molec cm}^{-3}$  (see Table S1 in the Supplement), with a multi-model mean of  $10.8 \pm 0.6 \times 10^5 \text{ molec cm}^{-3}$ , which is 10% higher than the estimate of Bousquet et al. (2005). Bousquet et al. (2005) showed that OH was susceptible to inter-annual fluctuations in concentrations of around  $8.5 \pm 1\%$ . The POLMIP multi-model mean OH concentration lies slightly outside this range of inter-annual variability, suggesting OH in the POLMIP models may be slightly overestimated. For comparison, multi-model mean OH concentrations of  $11.7 \pm 1.0 \times 10^5 \text{ molec cm}^{-3}$  (Voulgarakis et al., 2013) and  $11.1 \pm 1.7 \times 10^5 \text{ molec cm}^{-3}$  (Shindell et al., 2006) have been found previously, suggesting OH is also slightly overestimated in other models, assuming estimates of OH from methyl chloroform are correct.

## 5.2 Model differences in tracer transport

Even though differences exist in the model OH, it is reasonable to expect that some of the model spread in CO is explained by differences in simulated transport. Fig. 13 shows the seasonal zonal mean difference between the sum of the simulated 25-day fixed-loss regional tracer at two levels, 700 and 900 hPa. Negative values show there is more tracer in the upper 700 hPa level, positive values show that there is more tracer in the lower 900 hPa level and near zero values show a vertically well-mixed column (represented by the dashed line).

The largest positive values are found in the northern extra-tropics/lower mid-latitudes with very low CO differences in the Arctic and the tropics. In the Arctic, the models show near-zero differences in the winter and then negative differences in all other seasons. In winter, the so called “polar dome” extends quite far south due to strong surface cooling over land, allowing direct transport from Northern Eurasia into the Arctic at low levels whilst limiting transport from source regions that are warmer and more humid (e.g. East and South Asia) (e.g. Carlson, 1981). Due to surface heating, the “polar dome” begins to retract and move polewards as the year progresses, allowing emissions from more southerly located sources to have a larger impact (Law and Stohl, 2007; Law et al., 2014). As these emissions tend to be transported at higher altitudes (Stohl, 2006), they have a larger impact in the free troposphere compared to the surface explaining why the tracer concentrations are larger at 700 hPa compared to 900 hPa. The models all agree on the overall distribution and seasonal changes, however, there are some clear discrepancies in the magnitude of the tracer concentrations between models. These differences could be caused by a variety of model differences that can influence the transport of tracers, such as large-scale advection schemes, parameterised convection and the vertical and horizontal resolution of each model. For the models shown in Fig. 13, LMDZ and TOMCAT have a lower number of vertical levels below 700 hPa (9 and 8, respectively) compared to the other models (which have either 15 or 16 levels). However, they have very different vertical differences in the 25-day fixed lifetime tracer suggesting other processes apart from resolution are important.

It is likely that these vertical transport differences will play a role in the simulated Arctic concentration differences seen throughout Sect. 4.

### 5.3 Arctic sensitivities to regional anthropogenic and biomass burning emissions

The idealised tracers are used to compare the anthropogenic and biomass burning contributions from the three different source regions to the Arctic throughout the year. This is the first time biomass burning contributions have been summarised in this way. However, it is possible to compare the anthropogenic contributions found in this study to those in previous studies.

Figs. 14 and 15 show the absolute and fractional contributions from the anthropogenic and fire sourced regional tracers to the total tracer concentration in the Arctic. Fig. 14 shows the tracers with fixed 25-day loss rates and model-dependent transport, whilst Fig. 15 shows the tracers with varying OH-loss rates but fixed model transport. The tracers from each model are averaged seasonally into three different altitude bins over the Arctic region (north of 66° N).

The 25-day fixed-loss tracer highlights the seasonal changes in transport efficiencies from anthropogenic source regions (as the anthropogenic emissions are fixed throughout the year), but also shows the seasonality in biomass burning emissions. In winter, the highest concentrations of the 25-day fixed-loss tracer are located in the lowest altitude bin, with a clear decrease in absolute concentrations as altitude increases. This is in line with previous studies, which have shown poleward transport to occur mostly at low levels during winter (Klonecki et al., 2003; Stohl, 2006). As the year progresses, there is a clear shift in the concentrations to the higher altitude bins as transport of emissions at higher altitudes becomes more important in spring, summer and autumn (Klonecki et al., 2003).

Out of the three regional tracers, the European tracer shows the largest seasonal shift in transport efficiency to the Arctic. This is due to a large seasonal cycle in pollution export pathways from Europe (Duncan and Bey, 2004). The North American and Asian tracers show a much more consistent contribution to the Arctic total tracer concentration through-

out the year and troposphere. Europe is the most important anthropogenic source region at the surface in winter, with Asian emissions dominating at the highest altitudes, which is consistent with the concept of the “polar dome”. Similar to Asia, the contribution from North America is larger at higher altitudes, although the magnitude is much smaller. This is most likely due to a combination of lower emissions and different export efficiencies (Stohl, 2006). These are in broad agreement with multi-model idealised tracer results from Shindell et al. (2008), despite differences in emissions, tracer lifetimes and the area of the regions used. Results at the surface and in the mid-troposphere also agree with idealised tracer experiments performed by Klonecki et al. (2003), but differ in the upper troposphere where they showed Asian emissions to have the smallest contribution and North American to have a much larger fractional contribution. Increases in Asian emissions (Ohara et al., 2007) and decreases in European and North American emissions since 1990 (Duncan and Logan, 2008), which will be important in terms of the relative contributions, will be accounted for in the Streets v1.2 emissions inventory used for the POLMIP simulations. Klonecki et al. (2003) used the EDGAR v2.0 CO emission estimates for the year 1990, explaining why the Asian influence is lower than North American in their study.

In spring and summer, European and North American tracer concentrations decrease whilst Asian tracer concentrations marginally increase. This is in line with less efficient poleward transport from lower latitudes and the northward shift of the “polar dome” allowing more influence from Asia (Law et al., 2014). In autumn, a shift back towards winter transport conditions is visible.

In winter, there is very little contribution to the total tracer from biomass burning emissions, however, in spring, summer and autumn, biomass burning accounts for around 33, 43 and 19% of the total tracer, respectively (see Supplement Table S2). In spring, the largest overall contributions are found in the highest altitude bin, with the lower and middle bins showing larger overall contributions in summer and autumn. Asian fire emissions are clearly the largest source of boreal biomass burning tracer in the Arctic, with the largest contributions in spring and summer. North American and European fire emissions have their largest contributions in summer and Autumn, respectively. As previously mentioned,

spring burning began earlier than usual in 2008 and total 2008 boreal fire emissions were 21 % higher than average (Law et al., 2014). As the Arctic is highly sensitive to fire emission variability in the boreal regions (Monks et al., 2012), the fire contributions will be dependent on the year.

The regional tracers with OH-loss rates (Fig. 15) incorporate both seasonality in the CO lifetime due to changes in OH and seasonal changes in transport. Concentrations of the OH-loss tracers are at a maximum in winter when OH concentrations are low and poleward transport is efficient and at a minimum in summer when OH concentrations are high and transport is less efficient. The coincident changes in transport and tracer lifetime results in a seasonal change in absolute concentrations that is much larger than seen in the 25-day fixed-loss tracers. It should also be noted that inter-model differences in tracer concentrations are also larger for the OH-loss tracers compared to the 25-day fixed-loss tracers, however, the fractional contributions and therefore the relative importance of the different regions do not change by a large amount. This shows that the relative importance of the different source regions are robust across the models and will likely hold for “real” CO, even though individual model OH and transport differ. This therefore suggests that differences in model emission inventories are the most likely cause of differences in the relative importance of anthropogenic emission regions in Arctic source contribution studies for the same chemical species (as found here for Asian emissions in the Arctic upper troposphere in comparison to Klonecki et al., 2003).

#### **5.4 Inter-model variability in the Arctic**

In this section, inter-model variability in CO and O<sub>3</sub> at Arctic surface stations is discussed whilst the possible causes of variability are explored with the use of the idealised tracers to compare the impact of inter-model differences in transport and OH.

### 5.4.1 Model variability in carbon monoxide and ozone

The interquartile range (IQR) of simulated CO and O<sub>3</sub> at selected Arctic stations are shown in Fig. 4 (Sect. 4.1), both as an absolute value (in ppbv) and as a percent (of the seasonal mean observed concentration). The IQR gives a measure of the amount of spread in the POLMIP models.

For CO, the largest model spread occurs during autumn, when the multi-model mean bias is at a minimum. The lowest model spread occurs during spring, when the multi-model mean bias is high. The magnitude of the IQR is relatively similar across the two stations, suggesting the cause of the model spread may equally affect CO throughout the Arctic. The IQR of modelled O<sub>3</sub> shows a large amount of spread during winter and spring and the lowest amount of model spread during summer. Model spread is larger at Barrow, suggesting that the causes of the model spread do not equally impact the different station locations. In addition, the seasonality in model spread differs between O<sub>3</sub> and CO, suggesting the underlying causes of model spread may depend on the trace gas considered.

### 5.4.2 Model variability in OH and transport at Barrow

The 25-day fixed-loss anthropogenic regional tracers are used to identify model differences in regional export to the Arctic surface and the OH-loss anthropogenic regional tracers can be used as a proxy for differences in model chemistry that would result in different CO lifetimes and concentrations. Fig. 16 shows seasonal box and whisker plots at Barrow for the fixed-loss and OH-loss tracers from eight of the POLMIP models. The interquartile range (IQR) of each sample is shown both as an absolute concentration and as a percentage of the multi-model mean concentration.

The IQR of the fixed-loss tracers show that in winter the largest spread in model concentrations occurs in the European tracer. This coincides with the season when low-level export from Europe to the Arctic is particularly efficient (Duncan and Bey, 2004). This is shown by the winter European tracer having a higher median concentration than any other region and any other season (as expected from Sect. 5.3). In comparison to this, the model

spread in the European tracer is relatively low throughout the rest of the year. The Asian tracer has the second largest wintertime spread compared to the other regional tracers, however, the largest overall spread in this tracer is seen in summer, with model spread also being relatively high in spring and autumn. The North American tracer shows very little spread throughout the year suggesting simulated transport from North America to Barrow is relatively similar between models. The large variability in the Asian tracer in summer and the European tracer in winter is likely to explain some of the model spread seen in the Arctic CO and O<sub>3</sub> concentrations at Barrow shown in Fig. 4. Better constraints on simulated wintertime transport from Europe and transport from Asia throughout the year may help to reduce model spread in CO and O<sub>3</sub> in the Arctic.

The IQR of the OH-loss tracers shows that model spread is relatively consistent between the three different regions indicating that model OH differences affect all regions similarly, in contrast to inter-model transport differences. In terms of absolute concentrations, winter shows the largest overall spread in all three tracers. Spring and autumn have a similar amount of spread, with summer showing the lowest IQR. The percentage IQR is heavily weighted by the concentrations so that in summer when concentrations are very low, the percent IQR is very high ( $\geq 67\%$ ) and in winter when the concentrations are at a maximum the percent IQR is much lower ( $\leq 38\%$ ). However, even though this is the case, the percent IQR for the OH-loss tracers ( $25\% \geq \text{IQR} < 86\%$ ) is always larger than the respective percent IQR calculated for the 25-day fixed loss tracers ( $3\% < \text{IQR} < 40\%$ ). This shows that variations in OH are the dominant driver of inter-model differences in the total tracer concentrations at Barrow and that inter-model differences in the concentrations of pollutants transported from lower latitudes to the Arctic surface may be highly sensitive to inter-model OH differences if they react with OH.

## 5.5 Model variability throughout the Arctic

Model variability throughout the whole Arctic troposphere is considered using the coefficients of variation ( $\frac{\sigma}{\mu}$ ) for the OH-loss and 25-day fixed-loss tracers (see Fig. 17). This gives



a measure of inter-model variability regardless of the absolute concentrations by weighting the multi-model standard deviation by the multi-model mean concentration.

As found at Barrow, higher coefficients of variation, and therefore larger inter-model variability, are found for the OH-loss tracers compared to the 25-day fixed-loss tracers. Negative correlations between the burden of CO and OH concentrations averaged over the global troposphere ( $r^2 = 0.89$ ) and averaged over the Arctic region ( $r^2 = 0.30$ ) show that models with higher global mean OH have lower CO burdens, as expected (see Supplement Table S1). This relationship is much stronger on the global-scale compared to the Arctic alone, suggesting that the inter-model variability in the OH-loss tracers in the Arctic is mostly driven by inter-model OH differences at lower latitudes and not inter-model OH differences in the Arctic. This suggests that inter-model differences in CO seen throughout the Arctic troposphere are highly sensitive to inter-model differences in OH, when the same emissions are used. As shown in Fig. 12, inter-model OH differences are highly correlated with differences in model water vapour and photolysis rates, therefore improvements to these variables may reduce inter-model differences in Arctic trace species.

It is important to note that the results shown here are for idealised tracers and that the results may not be directly equated to simulations of CO and O<sub>3</sub>, where complex chemistry plays a role. For example, if a model has higher OH than other models, then it is likely to have a faster CO loss rate and therefore lower CO concentrations. However, it will also have faster oxidation of methane and other hydrocarbons and therefore more secondary production of CO to offset the higher loss rate of CO. The extent that these two opposing factors offset each other will be dependent on the model chemistry scheme. Shindell et al. (2008) concluded that oxidation rates, inferred from correlations between Arctic sensitivities and global CO lifetimes, did cause some inter-model differences in Arctic CO, but this was limited to the upper troposphere. The lack of any correlations in the lower troposphere is likely explained by the opposing impacts of OH on CO loss and secondary CO production, however, OH variability will still be particularly important when considering the production and loss terms of many reactive species in the Arctic.

## 6 Conclusions

We have used a range of surface, satellite and aircraft observations to evaluate multi-model simulations of CO and O<sub>3</sub> in the Arctic and sub-Arctic. The models include the same prescribed emissions for anthropogenic and biomass burning sources, removing one source of inter-model variability identified by previous model intercomparisons (Shindell et al., 2008), allowing the impacts of chemistry and transport differences on Arctic CO and O<sub>3</sub> burdens to be isolated.

The models broadly capture the observed seasonality of CO at the Arctic surface and over the mid-latitude lower troposphere. In agreement with previous studies, models generally underestimate CO in the Arctic at the surface, with the largest biases found during winter and spring. Our analysis, using POLARCAT 2008 spring and summer aircraft data, shows that these negative model CO biases also exist throughout the depth of the Arctic troposphere over Alaska in spring and several models remain biased low in the summer at high latitudes over Greenland and Canada.

MOPITT CO comparisons show that the models underestimate CO throughout the Northern Hemisphere, and that the large winter/spring bias in models may be shifted to higher altitudes in summer, when vertical transport by convection is important. These comparisons also show that model performance is not substantially better or worse over any of the mid-latitude source regions in contrast to MOZAIC aircraft observations, which show a slightly higher model CO bias over Europe compared to North America.

Simulated global mean tropospheric OH concentrations (multi-model mean =  $10.8 \pm 0.6 \times 10^5$  molecules cm<sup>-3</sup>) are slightly higher than estimates constrained by observations of methyl chloroform (Prinn et al., 2001; Bousquet et al., 2005), suggesting that the pervasive low biased CO concentrations in the Arctic may be slightly improved by better constraining simulated global mean OH concentrations, however, it is unlikely to explain the full extent of the bias. Evaluation of regional OH concentrations in the Arctic troposphere using ARCTAS aircraft observations demonstrates the highly variable model OH abundances. Models that have lower Arctic OH along the ARCTAS flight paths show no substantial improvement

in their negative CO biases, suggesting that there is little relationship between Arctic OH abundance and Arctic CO burden. This is consistent with OH-driven loss and secondary production of CO being largely controlled by OH abundances at lower latitudes, where OH concentrations are larger. In support of this, a significant correlation between modelled global CO burdens and global mean tropospheric OH concentrations is found ( $r^2 = 0.89$ ), whilst no significant correlation is found between CO burdens and mean OH concentrations in the Arctic alone ( $r^2 = 0.30$ ).

Previous studies have suggested emissions and missing chemical processes may explain the negative winter/spring bias in Arctic and Northern Hemispheric simulated CO (Kopacz et al., 2010; Mao et al., 2013). The GEOS-Chem model used in this study includes a new transition metal catalysed HO<sub>2</sub> uptake onto aerosol, that acts as a loss of HO<sub>2</sub> from the atmosphere rather than forming H<sub>2</sub>O<sub>2</sub>, as implemented in many other models. This increases the CO lifetime through loss of HO<sub>x</sub> (Mao et al., 2013) and likely explains the better agreement of GEOS-Chem with Arctic CO springtime observations in comparison to the other POLMIP models. However, summertime POLARCAT comparisons over Greenland and at surface stations suggest that this model overestimates CO in the lower troposphere in contrast to the other models. It is difficult in this multi-model assessment to isolate the impacts of the implementation of this HO<sub>2</sub> loss process in the GEOS-Chem model, and there is large uncertainty in appropriate values of uptake coefficients for this process on different aerosol types (Macintyre and Evans, 2011). The results here suggest that CO and oxidant budgets at high latitudes may be particularly sensitive to this process, and more extensive research into this process in Arctic simulations is warranted.

Model simulations of O<sub>3</sub> at Arctic surface stations show a much larger range in skill compared to those for CO. Models tend to underestimate Arctic O<sub>3</sub> concentrations at the surface throughout the year, with the exception of springtime at Barrow, when O<sub>3</sub> is overestimated due to a lack of halogen chemistry in most of the models. In addition, model O<sub>3</sub> is shown to be biased low against POLARCAT aircraft observations over Greenland in summer, predominantly near the surface and in the upper troposphere. This suggests that models may underestimate the influence of stratospheric O<sub>3</sub> in the Arctic upper tro-

posphere, and have problems simulating boundary layer processes controlling ozone over Greenland. In contrast, there is evidence that models overestimate MOZAIC observed  $O_3$  over Europe, particularly in the upper troposphere. This suggests a possible overestimation of stratospheric  $O_3$  influence at lower latitudes. Improvements to simulated boundary layer processes over snow/ice covered surfaces that affect  $O_3$  and stratospheric-tropospheric exchange may lead to improvement in model simulations of Arctic  $O_3$ .

Observations during summer demonstrate extensive perturbation of the Arctic troposphere by anthropogenic and fire influenced plumes, with associated enhancements in CO (Law et al. (2014) and references therein). The vertical structure of these enhancements is captured well by the POLMIP models, however the CO concentrations within the plumes are highly variable amongst models and often underestimated.  $O_3$  perturbations associated with the plumes also show large inter-model variability, which may point to different  $O_3$  production efficiencies in models. This could also be explained by deficiencies in model transport of mid-latitude sourced air masses over long distances into the Arctic in coarse global models, where plumes may become overly diffusive. This may also play a role in the negative bias in modelled CO within the same plumes.

As well as systematic model errors, the comparisons show a large amount of model spread in the concentrations of CO,  $O_3$  and OH, in agreement with previous multi-model comparisons (Shindell et al., 2008). This has implications for our confidence in the ability of models to accurately simulate atmospheric composition and climate responses to changes in mid-latitude emissions. We have used idealised CO-like tracers with either fixed-loss (25-days) and model dependent transport or model dependent OH-loss and the same transport (run within one model with loss calculated from the each model's monthly mean OH concentration fields from the full chemistry simulations) to compare the impacts of inter-model differences in OH and transport on Arctic concentrations of pollutants transported from mid-latitudes.

The fixed lifetime tracers show that model differences in transport from anthropogenic sources in Europe, during winter, and in Asia, throughout the year, are important sources of model variability at Barrow. Inter-model variability in OH leads to a similar amount of

model spread in the different regional anthropogenic OH-loss tracers showing OH differences equally affect all Arctic source regions. Comparisons of the OH-loss and 25-day fixed-loss tracers show that OH differences are a much larger source of inter-model variability in the tracer concentrations throughout the Arctic troposphere compared to transport differences. This demonstrates the importance of inter-model differences in OH for simulating tropospheric concentrations of reactive Arctic trace gases imported from the mid-latitudes. We propose that a better understanding of the inter-model differences in OH would help to understand and reduce uncertainty in model simulations of trace gas Arctic burdens. As suggested by the global and Arctic correlations between CO and OH, the inter-model variability in the Arctic concentrations of OH-loss CO-like tracers is mostly driven by inter-model OH differences at lower latitudes, not OH differences in the Arctic.

We have investigated possible drivers of variability in mean tropospheric OH for a subset of POLMIP models. Mean tropospheric OH was found to be significantly correlated with mean tropospheric water vapour concentrations, both in the Arctic ( $r^2 = 0.85$ ) and globally ( $r^2 = 0.91$ ). This suggests better constraints on water vapour may reduce inter-model variability in global mean OH concentrations and therefore Arctic CO. Mean OH concentrations and  $J(\text{O}^1\text{D})$  photolysis rates multiplied by  $\text{O}_3$  concentrations were found to be significantly correlated within the Arctic ( $r^2 = 0.65$ ), but not globally ( $r^2 = 0.01$ ). This is not in agreement with results from the ACCMIP model intercomparison study, which found a significant correlation between present-day OH and  $J(\text{O}^1\text{D})$  in a group of chemistry-climate models (Voulgarakis et al., 2013). Some of the POLMIP models have similar chemical mechanisms which may be affecting these correlations and therefore a more detailed study of the causes of inter-model OH differences would be beneficial.

Whilst inter-model differences in transport and, most notably, OH are shown to be important in terms of the inter-model differences in the absolute concentrations of the CO-like tracers in the Arctic, the fractional contributions and, therefore, the relative importance of the different source regions (North America, Europe and Asia) and different source types (anthropogenic and biomass burning) are similar for both the fixed-loss and OH-loss tracers. This suggests that differences in model emission inventories are the most likely cause of

differences in the relative importance of different anthropogenic emission regions in Arctic source contribution studies for the same chemical species. In support of this, Klonecki et al. (2003), found Asian anthropogenic emissions to have the smallest fractional contribution in the Arctic upper troposphere using emission estimates based on the year 1990, whereas, in our analysis, using more recent emissions estimates, Asian emissions dominate. This illustrates the potential impact of increasing Asian emissions on the Arctic over the 20 year period since 1990. The anthropogenic emissions used for the POLMIP model simulations result in similar conclusions in terms of the relative importance of different anthropogenic source regions to emissions sensitivities ( $\text{ppbv}(\text{CO}) \text{Tg}^{-1}$  emitted) to those reported by Shindell et al. (2008). Specifically, European emissions are most important at the surface in winter and Asian and North American emissions are most important at higher altitudes. In this study, emissions from fires in the boreal regions were also considered, and we showed that boreal fires can contribute 33, 43 and 19% to the total tracer in the Arctic in spring, summer and autumn, respectively, demonstrating the importance of fires as a source of Arctic pollution.

**The Supplement related to this article is available online at  
doi:10.5194/acpd-0-1-2015-supplement.**

*Acknowledgements.* This work was supported by the EurEX project, funded by the UK Natural Environmental Research Council (ref: NE/H020241/1). S. A. Monks and S. R. Arnold acknowledge support from the European Commission via the FP7 EUFAR Integrating Activity. The National Center for Atmospheric Research is funded by the U.S. National Science Foundation and operated by the University Corporation of Atmospheric Research. Author L. K. Emmons acknowledges support from the National Aeronautics and Space Administration under award no. NNX08AD22G issued through the Science Mission Directorate, Tropospheric Composition Program. Authors (K. S. Law, G. Ancellet, J. L. Thomas, J.-C. Raut, S. Turquety and Y. Long) acknowledge support from projects Agence National de Recherche (ANR) Climate Impact of Short-lived Climate Forcers and Methane in the Arctic (CLIMSLIP) Blanc SIMI 5–6 021 01 and CLIMSLIP-LEFE (CNRS-INSU). Valuable help with WRF-Chem simulations from T. Onishi (LATMOS/IPSL) and G. Pfister (NCAR) and with TOMCAT

maintenance from W. Feng (Leeds). J. Mao acknowledges the NOAA Climate Program Office grant NA13OAR4310071. Contributions by SMHI were funded by the Swedish Environmental Protection Agency under contract NV-09414-12 and through the Swedish Climate and Clean Air research program, SCAC. We thank the POLARCAT aircraft teams, especially the NASA ARCTAS, DLR-GRACE, and French ATR-42 teams. French ATR-42 campaigns and data analysis were part of POLARCAT-France, funded by French Agence Nationale de la Recherche (ANR), CNES, CNRS-INSU-LEFE, IPEV and EUFAR. Thanks is given to all those involved in the collection and provision of data, specifically NOAA/ESRL/WDCGG for the surface data, the MOZAIC- IAGOS project for aircraft data and the MOPITT team for satellite data.

## References

- ACIA: Arctic Climate Impact Assessment – Scientific Report, 21–60, Cambridge University Press, New York, USA, 2005.
- Alvarado, M. J., Logan, J. A., Mao, J., Apel, E., Riemer, D., Blake, D., Cohen, R. C., Min, K.-E., Perring, A. E., Browne, E. C., Wooldridge, P. J., Diskin, G. S., Sachse, G. W., Fuelberg, H., Sessions, W. R., Harrigan, D. L., Huey, G., Liao, J., Case-Hanks, A., Jimenez, J. L., Cubison, M. J., Vay, S. A., Weinheimer, A. J., Knapp, D. J., Montzka, D. D., Flocke, F. M., Pollock, I. B., Wennberg, P. O., Kurten, A., Crouse, J., Clair, J. M. St., Wisthaler, A., Mikoviny, T., Yantosca, R. M., Carouge, C. C., and Le Sager, P.: Nitrogen oxides and PAN in plumes from boreal fires during ARCTAS-B and their impact on ozone: an integrated analysis of aircraft and satellite observations, *Atmos. Chem. Phys.*, 10, 9739–9760, doi:10.5194/acp-10-9739-2010, 2010.
- Ancellet, G., Leclair de Bellevue, J., Mari, C., Nedelec, P., Kukui, A., Borbon, A., and Perros, P.: Effects of regional-scale and convective transports on tropospheric ozone chemistry revealed by aircraft observations during the wet season of the AMMA campaign, *Atmos. Chem. Phys.*, 9, 383–411, doi:10.5194/acp-9-383-2009, 2009.
- Andersson, C., Langner, J., and Bergström, R.: Interannual variation and trends in air pollution over Europe due to climate variability during 1958–2001 simulated with a regional CTM coupled to the ERA40 reanalysis, *Tellus B*, 59, 77–98, 2007.
- Archibald, A. T., Jenkin, M. E., and Shallcross, D. E.: An isoprene mechanism intercomparison, *Atmos. Environ.*, 44, 5356–5364, doi:10.1016/j.atmosenv.2009.09.016, *Atmospheric Chemical Mechanisms: Selected Papers from the 2008 Conference*, 2010.

- Arnold, S. R., Emmons, L. K., Monks, S. A., Law, K. S., Ridley, D. A., Turquety, S., Tilmes, S., Thomas, J. L., Bouarar, I., Flemming, J., Huijnen, V., Mao, J., Duncan, B. N., Steenrod, S., Yoshida, Y., Langner, J., and Long, Y.: Biomass burning influence on high latitude tropospheric ozone and reactive nitrogen in summer 2008: a multi-model analysis based on POLMIP simulations, *Atmos. Chem. Phys. Discuss.*, 14, 24573–24621, doi:10.5194/acpd-14-24573-2014, 2014.
- Atlas, E. L., Ridley, B. A., and Cantrell, C.: The Tropospheric Ozone Production about the Spring Equinox (TOPSE) Experiment: Introduction, *J. Geophys. Res.*, 108, 8353, doi:10.1029/2002JD003172, 2003.
- Barrie, L. A.: Arctic air pollution: an overview of current knowledge, *Atmos. Environ.* (1967), 20, 643–663, 1986.
- Barrie, L. A., Bottenheim, J. W., Schnell, R. C., Crutzen, P. J., and Rasmussen, R. A.: Ozone destruction and photochemical reactions at polar sunrise in the lower Arctic atmosphere, *Nature*, 334, 138–141, 1988.
- Barrie, L. A., Olson, M. P., and Oikawa, K. K.: The flux of anthropogenic sulphur into the arctic from mid-latitudes in 1979/80, *Atmos. Environ.* (1967), 23, 2505–2512, 1989.
- Bian, H., Colarco, P. R., Chin, M., Chen, G., Rodriguez, J. M., Liang, Q., Blake, D., Chu, D. A., da Silva, A., Darmenov, A. S., Diskin, G., Fuelberg, H. E., Huey, G., Kondo, Y., Nielsen, J. E., Pan, X., and Wisthaler, A.: Source attributions of pollution to the Western Arctic during the NASA ARCTAS field campaign, *Atmos. Chem. Phys.*, 13, 4707–4721, doi:10.5194/acp-13-4707-2013, 2013.
- Bottenheim, J. W. and Gallant, A. J.: Pan over the Arctic; observations during AGASP-2 in April 1986, *J. Atmos. Chem.*, 9, 301–316, doi:10.1007/BF00052839, 1989.
- Bottenheim, J. W., Gallant, A. G., and Brice, K. A.: Measurements of NO<sub>y</sub> species and O<sub>3</sub> at 82° N latitude, *Geophys. Res. Lett.*, 13, 113–116, 1986.
- Bourgeois, Q. and Bey, I.: Pollution transport efficiency toward the Arctic: sensitivity to aerosol scavenging and source regions, *J. Geophys. Res.-Atmos.*, 116, D08213, doi:10.1029/2010JD015096, 2011.
- Bousquet, P., Hauglustaine, D. A., Peylin, P., Carouge, C., and Ciais, P.: Two decades of OH variability as inferred by an inversion of atmospheric transport and chemistry of methyl chloroform, *Atmos. Chem. Phys.*, 5, 2635–2656, doi:10.5194/acp-5-2635-2005, 2005.
- Brock, C., Radke, L., Lyons, J., and Hobbs, P.: Arctic hazes in summer over Greenland and the North American Arctic. I: Incidence and origins, *J. Atmos. Chem.*, 9, 129–148, doi:10.1007/BF00052828, 1989.



- Brock, C. A., Cozic, J., Bahreini, R., Froyd, K. D., Middlebrook, A. M., McComiskey, A., Brioude, J., Cooper, O. R., Stohl, A., Aikin, K. C., de Gouw, J. A., Fahey, D. W., Ferrare, R. A., Gao, R.-S., Gore, W., Holloway, J. S., Hübler, G., Jefferson, A., Lack, D. A., Lance, S., Moore, R. H., Murphy, D. M., Nenes, A., Novelli, P. C., Nowak, J. B., Ogren, J. A., Peischl, J., Pierce, R. B., Pilewskie, P., Quinn, P. K., Ryerson, T. B., Schmidt, K. S., Schwarz, J. P., Sodemann, H., Spackman, J. R., Stark, H., Thomson, D. S., Thornberry, T., Veres, P., Watts, L. A., Warneke, C., and Wollny, A. G.: Characteristics, sources, and transport of aerosols measured in spring 2008 during the aerosol, radiation, and cloud processes affecting Arctic Climate (ARCPAC) Project, *Atmos. Chem. Phys.*, 11, 2423–2453, doi:10.5194/acp-11-2423-2011, 2011.
- Brune, W. H., Tan, D., Faloon, I. F., Jaeglés, L., Jacob, D. J., Heikes, B. G., Snow, J., Kondo, Y., Shetter, R., Sachse, G. W., Anderson, B., Gregory, G. L., Vay, S., Singh, H. B., Davis, D. D., Crawford, J. H., and Blake, D. R.: OH and HO<sub>2</sub> chemistry in the North Atlantic free troposphere, *Geophys. Res. Lett.*, 26, 3077–3080, doi:10.1029/1999GL900549, 1999.
- Carlson, T. N.: Speculations on the movement of polluted air to the Arctic, *Atmos. Environ.*, 15, 1473–1477, doi:10.1016/0004-6981(81)90354-1, Arctic Air Chemistry Proceedings of the Second Symposium, 1981.
- Chipperfield, M. P.: New version of the TOMCAT/SLIMCAT off-line chemical transport model: inter-comparison of stratospheric tracer experiments, *Q. J. Roy. Meteor. Soc.*, 132, 1179–1203, 2006.
- Deeter, M. N. and Martínez-Alonso, S. and Edwards, D. P. and Emmons, L. K. and Gille, J. C. and Worden, H. M. and Sweeney, C. and Pittman, J. V. and Daube, B. C. and Wofsy, S. C.: The MOPITT Version 6 product: Algorithm enhancements and validation., *Atmospheric Measurement Techniques*, 7, 3623-3632, doi:10.5194/amt-7-3623-2014, 2014.
- Deeter, M. N., Edwards, D. P., Gille, J. C., Emmons, L. K., Francis, G., Ho, S.-P., Mao, D., Masters, D., Worden, H., Drummond, J. R., and Novelli, P. C.: The MOPITT version 4 CO product: algorithm enhancements, validation, and long-term stability, *J. Geophys. Res.*, 115, D07306, doi:10.1029/2009JD013005, 2010.
- Deeter, M. N., Worden, H. M., Gille, J. C., Edwards, D. P., Mao, D., and Drummond, J. R.: MOPITT multispectral CO retrievals: origins and effects of geophysical radiance errors, *J. Geophys. Res.-Atmos.*, 116, D15303, doi:10.1029/2011JD015703, 2011.
- Donnell, E. A., Fish, D. J., Dicks, E. M., and Thorpe, A. J.: Mechanisms for pollutant transport between the boundary layer and the free troposphere, *J. Geophys. Res.-Atmos.*, 106, 7847–7856, doi:10.1029/2000JD900730, 2001.

- Duncan, B. N. and Bey, I.: A modeling study of the export pathways of pollution from Europe: seasonal and interannual variations (1987–1997), *J. Geophys. Res.*, 109, D08301, doi:10.1029/2003JD004079, 2004.
- Duncan, B. N. and Logan, J. A.: Model analysis of the factors regulating the trends and variability of carbon monoxide between 1988 and 1997, *Atmos. Chem. Phys.*, 8, 7389–7403, doi:10.5194/acp-8-7389-2008, 2008.
- Duncan, B. N., Strahan, S. E., Yoshida, Y., Steenrod, S. D., and Livesey, N.: Model study of the cross-tropopause transport of biomass burning pollution, *Atmos. Chem. Phys.*, 7, 3713–3736, doi:10.5194/acp-7-3713-2007, 2007.
- Eckhardt, S., Stohl, A., Wernli, H., James, P., Forster, C., and Spichtinger, N.: A 15-year climatology of warm conveyor belts, *J. Climate*, 17, 218–237, 2004.
- Emmons, L. K., Walters, S., Hess, P. G., Lamarque, J.-F., Pfister, G. G., Fillmore, D., Granier, C., Guenther, A., Kinnison, D., Laepple, T., Orlando, J., Tie, X., Tyndall, G., Wiedinmyer, C., Baughcum, S. L., and Kloster, S.: Description and evaluation of the Model for Ozone and Related chemical Tracers, version 4 (MOZART-4), *Geosci. Model Dev.*, 3, 43–67, doi:10.5194/gmd-3-43-2010, 2010.
- Emmons, L., Arnold, S. R., Monks, S., Huijnen, V., Tilmes, S., Law, K., Turquety, S., Long, Y., Thomas, J., Bouarar, I., Duncan, B., Steenrod, S., Strode, S., Flemming, J., Mao, J., and Langner, J.: The POLARCAT Model Intercomparison Project (POLMIP): overview and evaluation with observations, *Atmos. Chem. Phys. Discuss.*, in preparation, 2014.
- Fast, J. D., Gustafson, W. I., Easter, R. C., Zaveri, R. A., Barnard, J. C., Chapman, E. G., Grell, G. A., and Peckham, S. E.: Evolution of ozone, particulates, and aerosol direct radiative forcing in the vicinity of Houston using a fully coupled meteorology-chemistry-aerosol model, *J. Geophys. Res.-Atmos.*, 111, D21305, doi:10.1029/2005JD006721, 2006.
- Fisher, J. A., Jacob, D. J., Purdy, M. T., Kopacz, M., Le Sager, P., Carouge, C., Holmes, C. D., Yantosca, R. M., Batchelor, R. L., Strong, K., Diskin, G. S., Fuelberg, H. E., Holloway, J. S., Hyer, E. J., McMillan, W. W., Warner, J., Streets, D. G., Zhang, Q., Wang, Y., and Wu, S.: Source attribution and interannual variability of Arctic pollution in spring constrained by aircraft (ARCTAS, ARCPAC) and satellite (AIRS) observations of carbon monoxide, *Atmos. Chem. Phys.*, 10, 977–996, doi:10.5194/acp-10-977-2010, 2010.
- Flemming, J., Huijnen, V., Arteta, J., Bechtold, P., Beljaars, A., Blechschmidt, A.-M., Josse, B., Diamantakis, M., Engelen, R. J., Gaudel, A., Inness, A., Jones, L., Katragkou, E., Marecal, V., Peuch, V.-H., Richter, A., Schultz, M. G., Stein, O., and Tsikerdekis, A.: Tropospheric chemistry

- in the integrated forecasting system of ECMWF, *Geosci. Model Dev. Discuss.*, 7, 7733–7803, doi:10.5194/gmdd-7-7733-2014, 2014.
- Fuelberg, H. E., Harrigan, D. L., and Sessions, W.: A meteorological overview of the ARCTAS 2008 mission, *Atmos. Chem. Phys.*, 10, 817–842, doi:10.5194/acp-10-817-2010, 2010.
- Gerbig, C., Schmitgen, S., Kley, D., Volz-Thomas, A., Dewey, K., and Haaks, D.: An improved fast-response vacuum-UV resonance fluorescence CO instrument, *J. Geophys. Res.*, 104, 1699–1704, 1999.
- Grell, G. A., Peckham, S. E., Schmitz, R., McKeen, S. A., Frost, G., Skamarock, W. C., and Eder, B.: Fully coupled ‘online’ chemistry within the WRF model, *Atmos. Environ.*, 39, 6957–6975, doi:10.1016/j.atmosenv.2005.04.027, 2005.
- Guenther, A. B., Jiang, X., Heald, C. L., Sakulyanontvittaya, T., Duhl, T., Emmons, L. K., and Wang, X.: The Model of Emissions of Gases and Aerosols from Nature version 2.1 (MEGAN2.1): an extended and updated framework for modeling biogenic emissions, *Geosci. Model Dev.*, 5, 1471–1492, doi:10.5194/gmd-5-1471-2012, 2012.
- Harriss, R. C., Wofsy, S. C., Hoell, J. M., J., Bendura, R. J., Drewry, J. W., McNeal, R. J., Pierce, D., Rabine, V., and Snell, R. L.: The Arctic Boundary Layer Expedition (ABLE-3B): July–August 1990, *J. Geophys. Res.*, 99, 1635–1643, 1994.
- Helmig, D., Oltmans, S. J., Carlson, D., Lamarque, J.-F., Jones, A., Labuschagne, C., Anlauf, K., and Hayden, K.: A review of surface ozone in the polar regions, *Atmos. Environ.*, 41, 5138–5161, 2007.
- Helmig, D., Cohen, L. D., Bocquet, F., Oltmans, S., Grachev, A., and Neff, W.: Spring and summer-time diurnal surface ozone fluxes over the polar snow at Summit, Greenland, *Geophys. Res. Lett.*, 36, L08809, doi:10.1029/2008GL036549, 2009.
- Hirdman, D., Sodemann, H., Eckhardt, S., Burkhardt, J. F., Jefferson, A., Mefford, T., Quinn, P. K., Sharma, S., Ström, J., and Stohl, A.: Source identification of short-lived air pollutants in the Arctic using statistical analysis of measurement data and particle dispersion model output, *Atmos. Chem. Phys.*, 10, 669–693, doi:10.5194/acp-10-669-2010, 2010.
- Hourdin, F., Musat, I., Bony, S., Braconnot, P., Codron, F., Dufresne, J.-L., Fairhead, L., Filiberti, M.-A., Friedlingstein, P., Grandpeix, J.-Y., Krinner, G., LeVan, P., Li, Z.-X., and Lott, F.: The LMDZ4 general circulation model: climate performance and sensitivity to parametrized physics with emphasis on tropical convection, *Clim. Dynam.*, 27, 787–813, doi:10.1007/s00382-006-0158-0, 2006.

- Hov, Ø., and Flatøy, F.: Convective redistribution of ozone and oxides of nitrogen in the troposphere over Europe in summer and fall, *J. Atmos. Chem.*, 28, 319–337, doi:10.1023/A:1005780730600, 1997.
- Huijnen, V., Williams, J., van Weele, M., van Noije, T., Krol, M., Dentener, F., Segers, A., Houweling, S., Peters, W., de Laat, J., Boersma, F., Bergamaschi, P., van Velthoven, P., Le Sager, P., Eskes, H., Alkemade, F., Scheele, R., Nédélec, P., and Pätz, H.-W.: The global chemistry transport model TM5: description and evaluation of the tropospheric chemistry version 3.0, *Geosci. Model Dev.*, 3, 445–473, doi:10.5194/gmd-3-445-2010, 2010.
- Iversen, T.: Numerical modelling of the long range atmospheric transport of sulphur dioxide and particulate sulphate to the arctic, *Atmos. Environ.* (1967), 23, 2571–2595, doi:10.1016/0004-6981(89)90267-9, 1989.
- Jacob, D. J., Crawford, J. H., Maring, H., Clarke, A. D., Dibb, J. E., Emmons, L. K., Ferrare, R. A., Hostetler, C. A., Russell, P. B., Singh, H. B., Thompson, A. M., Shaw, G. E., McCauley, E., Pederson, J. R., and Fisher, J. A.: The Arctic Research of the Composition of the Troposphere from Aircraft and Satellites (ARCTAS) mission: design, execution, and first results, *Atmos. Chem. Phys.*, 10, 5191–5212, doi:10.5194/acp-10-5191-2010, 2010.
- Jaegle, L.: Pumping up surface air, *Science*, 315, 772–773, doi:10.1126/science.1138988, 2007.
- Khalil, M. A. K. and Rasmussen, R. A.: Statistical analysis of trace gases in Arctic haze, *Geophys. Res. Lett.*, 11, 437–440, 1984.
- Khalil, M. A. K., and Rasmussen, R. A.: Global decrease in atmospheric carbon monoxide concentration, *Nature*, 370, 639–641, 1994.
- Klonecki, A., Hess, P., Emmons, L., Smith, L., Orlando, J., and Blake, D.: Seasonal changes in the transport of pollutants into the Arctic troposphere—model study, *J. Geophys. Res.*, 108, 8367, doi:10.1029/2001JD001390, 2003.
- Koch, D. and Hansen, J.: Distant origins of Arctic black carbon: a Goddard institute for space studies modell experiment, *J. Geophys. Res.*, 110, D04204, doi:10.1029/2004JD005296, 2005.
- Kopacz, M., Jacob, D. J., Fisher, J. A., Logan, J. A., Zhang, L., Megretskaja, I. A., Yantosca, R. M., Singh, K., Henze, D. K., Burrows, J. P., Buchwitz, M., Khlystova, I., McMillan, W. W., Gille, J. C., Edwards, D. P., Eldering, A., Thouret, V., and Nedelec, P.: Global estimates of CO sources with high resolution by adjoint inversion of multiple satellite datasets (MOPITT, AIRS, SCIAMACHY, TES), *Atmos. Chem. Phys.*, 10, 855–876, doi:10.5194/acp-10-855-2010, 2010.
- Lamarque, J.-F., Emmons, L. K., Hess, P. G., Kinnison, D. E., Tilmes, S., Vitt, F., Heald, C. L., Holland, E. A., Lauritzen, P. H., Neu, J., Orlando, J. J., Rasch, P. J., and Tyndall, G. K.: CAM-chem:

- description and evaluation of interactive atmospheric chemistry in the Community Earth System Model, *Geosci. Model Dev.*, 5, 369–411, doi:10.5194/gmd-5-369-2012, 2012.
- Law, K. S. and Stohl, A.: Arctic air pollution: origins and impacts, *Science*, 315, 1537–1540, doi:10.1126/science.1137695, 2007.
- Law, K., Stohl, A., Quinn, P., Brock, C., Burkhardt, J., Paris, J., Ancellet, G., Singh, H., Roiger, A., Schlager, H., Dibb, J., Jacob, D., Arnold, S., Pelon, J., and Thomas, J.: Arctic air pollution: new insights from POLARCAT-IPY, *B. Am. Meteorol. Soc.*, doi:10.1175/BAMS-D-13-00017.1, in press, 2014.
- Lawrence, M. G., Jöckel, P., and von Kuhlmann, R.: What does the global mean OH concentration tell us?, *Atmos. Chem. Phys.*, 1, 37–49, doi:10.5194/acp-1-37-2001, 2001.
- Legrand, M., De Angelis, M., Staffelbach, T., Neftel, A., and Stauffer, B.: Large perturbations of ammonium and organic acids content in the summit Greenland Ice Core. Fingerprint from forest fires?, *Geophys. Res. Lett.*, 19, 473–475, 1992.
- Levy, H.: Normal atmosphere: large radical and formaldehyde concentrations predicted, *Science*, 173, 141–143, doi:10.1126/science.173.3992.141, 1971.
- Logan, J. A., Prather, M. J., Wofsy, S. C., and McElroy, M. B.: Tropospheric chemistry: a global perspective, *J. Geophys. Res.-Oceans*, 86, 7210–7254, doi:10.1029/JC086iC08p07210, 1981.
- Macintyre, H. L. and Evans, M. J.: Parameterisation and impact of aerosol uptake of HO<sub>2</sub> on a global tropospheric model, *Atmos. Chem. Phys. Discuss.*, 11, 16311–16334, doi:10.5194/acpd-11-16311-2011, 2011.
- Mao, J., Jacob, D. J., Evans, M. J., Olson, J. R., Ren, X., Brune, W. H., Clair, J. M. St., Crouse, J. D., Spencer, K. M., Beaver, M. R., Wennberg, P. O., Cubison, M. J., Jimenez, J. L., Fried, A., Weibring, P., Walega, J. G., Hall, S. R., Weinheimer, A. J., Cohen, R. C., Chen, G., Crawford, J. H., McNaughton, C., Clarke, A. D., Jaeglé, L., Fisher, J. A., Yantosca, R. M., Le Sager, P., and Carouge, C.: Chemistry of hydrogen oxide radicals (HO<sub>x</sub>) in the Arctic troposphere in spring, *Atmos. Chem. Phys.*, 10, 5823–5838, doi:10.5194/acp-10-5823-2010, 2010.
- Mao, J., Fan, S., Jacob, D. J., and Travis, K. R.: Radical loss in the atmosphere from Cu-Fe redox coupling in aerosols, *Atmos. Chem. Phys.*, 13, 509–519, doi:10.5194/acp-13-509-2013, 2013.
- Monks, S. A.: A model study of chemistry and transport in the Arctic troposphere, Ph. D. thesis, University of Leeds, 2011.
- Monks, S. A., Arnold, S. R., and Chipperfield, M. P.: Evidence for El Niño–Southern Oscillation (ENSO) influence on Arctic CO interannual variability through biomass burning emissions, *Geophys. Res. Lett.*, 39, L14804, doi:10.1029/2012GL052512, 2012.

- Murray, L. T., Mickley, L. J., Kaplan, J. O., Sofen, E. D., Pfeiffer, M., and Alexander, B.: Factors controlling variability in the oxidative capacity of the troposphere since the Last Glacial Maximum, *Atmos. Chem. Phys.*, 14, 3589–3622, doi:10.5194/acp-14-3589-2014, 2014.
- Nedelec, P., Cammas, J.-P., Thouret, V., Athier, G., Cousin, J.-M., Legrand, C., Abonnel, C., Lecoq, F., Cayez, G., and Marizy, C.: An improved infrared carbon monoxide analyser for routine measurements aboard commercial Airbus aircraft: technical validation and first scientific results of the MOZAIC III programme, *Atmos. Chem. Phys.*, 3, 1551–1564, doi:10.5194/acp-3-1551-2003, 2003.
- Novelli, P. C., Masarie, K. A., Tans, P. P., and Lang, P. M.: Recent changes in atmospheric carbon monoxide, *Science*, 263, 1587–1590, doi:10.1126/science.263.5153.1587, 1994.
- Novelli, P., Masarie, K., and Lang, P.: Distributions and recent changes of carbon monoxide in the lower troposphere, *J. Geophys. Res.*, 103, 19015–19033, 1998.
- Ohara, T., Akimoto, H., Kurokawa, J., Horii, N., Yamaji, K., Yan, X., and Hayasaka, T.: An Asian emission inventory of anthropogenic emission sources for the period 1980–2020, *Atmos. Chem. Phys.*, 7, 4419–4444, doi:10.5194/acp-7-4419-2007, 2007.
- Oltmans, S. J. and Levy, H.: Surface ozone measurements from a global network, *Atmos. Environ.*, 28, 9–24, doi:10.1016/1352-2310(94)90019-1, 1994.
- Oltmans, S., Lefohn, A., Harris, J., Galbally, I., Scheel, H., Bodeker, G., Brunke, E., Claude, H., Tarasick, D., Johnson, B., Simmonds, P., Shadwick, D., Anlauf, K., Hayden, K., Schmidlin, F., Fujimoto, T., Akagi, K., Meyer, C., Nichol, S., Davies, J., Redondas, A., and Cuevas, E.: Long-term changes in tropospheric ozone, *Atmos. Environ.*, 40, 3156–3173, doi:10.1016/j.atmosenv.2006.01.029, 2006.
- Oltmans, S. J., Lefohn, A. S., Scheel, H. E., Harris, J. M., Levy, H., I., Galbally, I. E., Brunke, E., Meyer, C. P., Lathrop, J. A., Johnson, B. J., Shadwick, D. S., Cuevas, E., Schmidlin, F. J., Tarasick, D. W., Claude, H., Kerr, J. B., Uchino, O., and Mohnen, V.: Trends of ozone in the troposphere, *Geophys. Res. Lett.*, 25, 139–142, 1998.
- Paris, J.-D., Stohl, A., Nédélec, P., Arshinov, M. Yu., Panchenko, M. V., Shmargunov, V. P., Law, K. S., Belan, B. D., and Ciais, P.: Wildfire smoke in the Siberian Arctic in summer: source characterization and plume evolution from airborne measurements, *Atmos. Chem. Phys.*, 9, 9315–9327, doi:10.5194/acp-9-9315-2009, 2009.
- Parrella, J. P., Jacob, D. J., Liang, Q., Zhang, Y., Mickley, L. J., Miller, B., Evans, M. J., Yang, X., Pyle, J. A., Theys, N., and Van Roozendael, M.: Tropospheric bromine chemistry: implica-

- tions for present and pre-industrial ozone and mercury, *Atmos. Chem. Phys.*, 12, 6723–6740, doi:10.5194/acp-12-6723-2012, 2012.
- Patra PK; Houweling S; Krol M; Bousquet P; Belikov D; Bergmann D; Bian H; Cameron-Smith P; Chipperfield MP; Corbin K; Fortems-Cheiney A; Fraser A; Gloor E; Hess P; Ito A; Kawa SR; Law RM; Loh Z; Maksyutov S; Meng L; Palmer PI; Prinn RG; Rigby M; Saito R; Wilson C: TransCom model simulations of CH<sub>4</sub> and related species: Linking transport, surface flux and chemical loss with CH<sub>4</sub> variability in the troposphere and lower stratosphere, *Atmos. Chem. Phys.*, 11, 12813–12837, doi:10.5194/acp-11-12813-2011, 2011.
- Pommier, M., Law, K. S., Clerbaux, C., Turquety, S., Hurtmans, D., Hadji-Lazarou, J., Coheur, P.-F., Schlager, H., Ancellet, G., Paris, J.-D., Nédélec, P., Diskin, G. S., Podolske, J. R., Holloway, J. S., and Bernath, P.: IASI carbon monoxide validation over the Arctic during POLARCAT spring and summer campaigns, *Atmos. Chem. Phys.*, 10, 10655–10678, doi:10.5194/acp-10-10655-2010, 2010.
- Prinn, R. G., Huang, J., Weiss, R. F., Cunnold, D. M., Fraser, P. J., Simmonds, P. G., McCulloch, A., Harth, C., Salameh, P., O'Doherty, S., Wang, R. H. J., Porter, L., and Miller, B. R.: Evidence for substantial variations of atmospheric hydroxyl radicals in the past two decades, *Science*, 292, 1882–1888, 2001.
- Quinn, P. K., Bates, T. S., Baum, E., Doubleday, N., Fiore, A. M., Flanner, M., Fridlind, A., Garrett, T. J., Koch, D., Menon, S., Shindell, D., Stohl, A., and Warren, S. G.: Short-lived pollutants in the Arctic: their climate impact and possible mitigation strategies, *Atmos. Chem. Phys.*, 8, 1723–1735, doi:10.5194/acp-8-1723-2008, 2008.
- Rahn, K. A.: Progress in Arctic air chemistry, 1980–1984, *Atmos. Environ.* (1967), 19, 1987–1994, doi:10.1016/0004-6981(85)90107-6, 1985.
- Randerson, J. T., Liu, H., Flanner, M. G., Chambers, S. D., Jin, Y., Hess, P. G., Pfister, G., Mack, M. C., Treseder, K. K., Welp, L. R., Chapin, F. S., Harden, J. W., Goulden, M. L., Lyons, E., Neff, J. C., Schuur, E. A. G., and Zender, C. S.: The impact of boreal forest fire on climate warming, *Science*, 314, 1130–1132, doi:10.1126/science.1132075, 2006.
- Robertson, L., Langner, J., and Engardt, M.: An Eulerian limited area atmospheric transport model, *J. Appl. Meteorol.*, 38, 90–210, 1999.
- Roiger, A., Schlager, H., Schäfler, A., Huntrieser, H., Scheibe, M., Aufmhoff, H., Cooper, O. R., Sodemann, H., Stohl, A., Burkhardt, J., Lazzara, M., Schiller, C., Law, K. S., and Arnold, F.: In-situ observation of Asian pollution transported into the Arctic lowermost stratosphere, *Atmos. Chem. Phys.*, 11, 10975–10994, doi:10.5194/acp-11-10975-2011, 2011.

- Sachse, G. W., and Hill, G. F.: Fast-response, high precision carbon monoxide sensor using a tunable diode laser absorption technique, *J. Geophys. Res.*, 92, 2071–2081, 1987.
- Schnell, R. C. and Raatz, W. E.: Vertical and horizontal characteristics of Arctic haze during AGASP: Alaskan Arctic, *Geophys. Res. Lett.*, 11, 369–372, 1984.
- Schnell, R. C., Watson, T. B., and Bodhaine, B. A.: NOAA WP-3D instrumentation and flight operations on AGASP-II, *J. Atmos. Chem.*, 9, 3–16, doi:10.1007/BF00052822, 1989.
- Serreze, M. C. and Francis, J. A.: The Arctic amplification debate, *Climatic Change*, 76, 241–264, 2006.
- Shaw, G. E.: The Arctic haze phenomenon, *B. Am. Meteorol. Soc.*, 76, 2403–2413, 1995.
- Shindell, D.: Local and remote contributions to Arctic warming, *Geophys. Res. Lett.*, 34, L14704, doi:10.1029/2007GL030221, 2007.
- Shindell, D. and Faluvegi, G.: Climate response to regional radiative forcing during the twentieth century, *Nat. Geosci.*, 2, 294–300, 2009.
- Shindell, D. T., Faluvegi, G., Stevenson, D. S., Krol, M. C., Emmons, L. K., Lamarque, J.-F., Pétron, G., Dentener, F. J., Ellingsen, K., Schultz, M. G., Wild, O., Amann, M., Atherton, C. S., Bergmann, D. J., Bey, I., Butler, T., Cofala, J., Collins, W. J., Derwent, R. G., Doherty, R. M., Drevet, J., Eskes, H. J., Fiore, A. M., Gauss, M., Hauglustaine, D. A., Horowitz, L. W., Isakson, I. S. A., Lawrence, M. G., Montanaro, V., Müller, J.-F., Pitari, G., Prather, M. J., Pyle, J. A., Rast, S., Rodriguez, J. M., Sanderson, M. G., Savage, N. H., Strahan, S. E., Sudo, K., Szopa, S., Unger, N., van Noije, T. P. C., and Zeng, G.: Multimodel simulations of carbon monoxide: comparison with observations and projected near-future changes, *J. Geophys. Res.*, 111, D19306, doi:10.1029/2006JD007100, 2006.
- Shindell, D. T., Chin, M., Dentener, F., Doherty, R. M., Faluvegi, G., Fiore, A. M., Hess, P., Koch, D. M., MacKenzie, I. A., Sanderson, M. G., Schultz, M. G., Schulz, M., Stevenson, D. S., Teich, H., Textor, C., Wild, O., Bergmann, D. J., Bey, I., Bian, H., Cuvelier, C., Duncan, B. N., Folberth, G., Horowitz, L. W., Jonson, J., Kaminski, J. W., Marmer, E., Park, R., Pringle, K. J., Schroeder, S., Szopa, S., Takemura, T., Zeng, G., Keating, T. J., and Zuber, A.: A multi-model assessment of pollution transport to the Arctic, *Atmos. Chem. Phys.*, 8, 5353–5372, doi:10.5194/acp-8-5353-2008, 2008.
- Sodemann, H., Pommier, M., Arnold, S. R., Monks, S. A., Stebel, K., Burkhardt, J. F., Hair, J. W., Diskin, G. S., Clerbaux, C., Coheur, P.-F., Hurtmans, D., Schlager, H., Blechschmidt, A.-M., Kristjánsson, J. E., and Stohl, A.: Episodes of cross-polar transport in the Arctic troposphere



- during July 2008 as seen from models, satellite, and aircraft observations, *Atmos. Chem. Phys.*, 11, 3631–3651, doi:10.5194/acp-11-3631-2011, 2011.
- Solberg, S., Dye, C., Schmidbauer, N., Herzog, A., and Gehrig, R.: Carbonyls and nonmethane hydrocarbons at rural European sites from the mediterranean to the arctic, *J. Atmos. Chem.*, 25, 33–66, doi:10.1007/BF00053285, 1996.
- Stohl, A.: A 1-year Lagrangian climatology of airstreams in the Northern Hemisphere troposphere and lowermost stratosphere, *J. Geophys. Res.*, 106, 7263–7279, 2001.
- Stohl, A.: Characteristics of atmospheric transport into the Arctic troposphere, *J. Geophys. Res.*, 111, D11306, doi:10.1029/2005JD006888, 2006.
- Strahan, S. E., Duncan, B. N., and Hoor, P.: Observationally derived transport diagnostics for the lowermost stratosphere and their application to the GMI chemistry and transport model, *Atmos. Chem. Phys.*, 7, 2435–2445, doi:10.5194/acp-7-2435-2007, 2007.
- Szopa, S., Balkanski, Y., Schulz, M., Bekki, S., Cugnet, D., Fortems-Cheiney, A., Turquety, S., Cozic, A., Déandreis, C., Hauglustaine, D., Idelkadi, A., Lathière, J., Lefevre, F., Marchand, M., Vuolo, R., Yan, N., and Dufresne, J.-L.: Aerosol and ozone changes as forcing for climate evolution between 1850 and 2100, *Clim. Dynam.*, 40, 2223–2250, doi:10.1007/s00382-012-1408-y, 2013.
- Thomas, J. L., Raut, J.-C., Law, K. S., Marelle, L., Ancellet, G., Ravetta, F., Fast, J. D., Pfister, G., Emmons, L. K., Diskin, G. S., Weinheimer, A., Roiger, A., and Schlager, H.: Pollution transport from North America to Greenland during summer 2008, *Atmos. Chem. Phys.*, 13, 3825–3848, doi:10.5194/acp-13-3825-2013, 2013.
- Thouret, V., Marengo, A., Logan, J. A., Nédélec, P., and Grouhel, C.: Comparisons of ozone measurements from the MOZAIC airborne program and the ozone sounding network at eight locations, *J. Geophys. Res.-Atmos.*, 103, 25695–25720, doi:10.1029/98JD02243, 1998.
- Tilmes, S., Emmons, L. K., Law, K. S., Ancellet, G., Schlager, H., Paris, J.-D., Fuelberg, H. E., Streets, D. G., Wiedinmyer, C., Diskin, G. S., Kondo, Y., Holloway, J., Schwarz, J. P., Spackman, J. R., Campos, T., Nédélec, P., and Panchenko, M. V.: Source contributions to Northern Hemisphere CO and black carbon during spring and summer 2008 from POLARCAT and START08/preHIPPO observations and MOZART-4, *Atmos. Chem. Phys. Discuss.*, 11, 5935–5983, doi:10.5194/acpd-11-5935-2011, 2011.
- Tilmes, S., Lamarque, J.-F., Emmons, L. K., Kinnison, D. E., Ma, P.-L., Liu, X., Ghan, S., Bardeen, C., Arnold, S., Deeter, M., Vitt, F., Ryerson, T., Elkins, J. W., Moore, F., and Spackman, R.: Description

- and evaluation of tropospheric chemistry and aerosols in the Community Earth System Model (CESM1.2), *Geosci. Model Dev. Discuss.*, 7, 8875-8940, doi:10.5194/gmdd-7-8875-2014, 2014.
- Voulgarakis, A., Naik, V., Lamarque, J.-F., Shindell, D. T., Young, P. J., Prather, M. J., Wild, O., Field, R. D., Bergmann, D., Cameron-Smith, P., Cionni, I., Collins, W. J., Dalsøren, S. B., Doherty, R. M., Eyring, V., Faluvegi, G., Folberth, G. A., Horowitz, L. W., Josse, B., MacKenzie, I. A., Nagashima, T., Plummer, D. A., Righi, M., Rumbold, S. T., Stevenson, D. S., Strode, S. A., Sudo, K., Szopa, S., and Zeng, G.: Analysis of present day and future OH and methane lifetime in the ACCMIP simulations, *Atmos. Chem. Phys.*, 13, 2563–2587, doi:10.5194/acp-13-2563-2013, 2013.
- Walker, T. W., Jones, D. B. A., Parrington, M., Henze, D. K., Murray, L. T., Bottenheim, J. W., Anlauf, K., Worden, J. R., Bowman, K. W., Shim, C., Singh, K., Kopacz, M., Tarasick, D. W., Davies, J., von der Gathen, P., Thompson, A. M., and Carouge, C. C.: Impacts of midlatitude precursor emissions and local photochemistry on ozone abundances in the Arctic, *J. Geophys. Res.-Atmos.*, 117, D01305, doi:10.1029/2011JD016370, 2012.
- Warneke, C., Bahreini, R., Brioude, J., Brock, C. A., de Gouw, J. A., Fahey, D. W., Froyd, K. D., Holloway, J. S., Middlebrook, A., Miller, L., Montzka, S., Murphy, D. M., Peischl, J., Ryerson, T. B., Schwarz, J. P., Spackman, J. R., and Veres, P.: Biomass burning in Siberia and Kazakhstan as an important source for haze over the Alaskan Arctic in April 2008, *Geophys. Res. Lett.*, 36, L02813, doi:10.1029/2008GL036194, 2009.
- Warneke, C., Froyd, K. D., Brioude, J., Bahreini, R., Brock, C. A., Cozic, J., de Gouw, J. A., Fahey, D. W., Ferrare, R., Holloway, J. S., Middlebrook, A. M., Miller, L., Montzka, S., Schwarz, J. P., Sodemann, H., Spackman, J. R., and Stohl, A.: An important contribution to springtime Arctic aerosol from biomass burning in Russia, *Geophys. Res. Lett.*, 37, L01801, doi:10.1029/2009GL041816, 2010.
- Weinheimer, A. J., Walega, J. G., Ridley, B. A., Gary, B. L., Blake, D. R., Blake, N. J., Rowland, F. S., Sachse, G. W., Anderson, B. E., and Collins, J. E.: Meridional distributions of NO<sub>x</sub>, NO<sub>y</sub>, and other species in the lower stratosphere and upper troposphere during AASE II, *Geophys. Res. Lett.*, 21, 2583–2586, 1994.
- Wespes, C., Emmons, L., Edwards, D. P., Hannigan, J., Hurtmans, D., Saunio, M., Coheur, P.-F., Clerbaux, C., Coffey, M. T., Batchelor, R. L., Lindenmaier, R., Strong, K., Weinheimer, A. J., Nowak, J. B., Ryerson, T. B., Crounse, J. D., and Wennberg, P. O.: Analysis of ozone and nitric acid in spring and summer Arctic pollution using aircraft, ground-based, satellite observations

- and MOZART-4 model: source attribution and partitioning, *Atmos. Chem. Phys.*, 12, 237–259, doi:10.5194/acp-12-237-2012, 2012.
- Wiedinmyer, C., Akagi, S. K., Yokelson, R. J., Emmons, L. K., Al-Saadi, J. A., Orlando, J. J., and Soja, A. J.: The Fire INventory from NCAR (FINN): a high resolution global model to estimate the emissions from open burning, *Geosci. Model Dev.*, 4, 625–641, doi:10.5194/gmd-4-625-2011, 2011.
- Wild, O. and Akimoto, H.: Intercontinental transport of ozone and its precursors in a three-dimensional global CTM, *J. Geophys. Res.-Atmos.*, 106, 27729–27744, doi:10.1029/2000JD000123, 2001.
- Williams, J. E., van Velthoven, P. F. J., and Brenninkmeijer, C. A. M.: Quantifying the uncertainty in simulating global tropospheric composition due to the variability in global emission estimates of Biogenic Volatile Organic Compounds, *Atmos. Chem. Phys.*, 13, 2857–2891, doi:10.5194/acp-13-2857-2013, 2013.
- Wofsy, S. C., Sachse, G. W., Gregory, G. L., Blake, D. R., Bradshaw, J. D., Sandholm, S. T., Singh, H. B., Barrick, J. A., Harriss, R. C., Talbot, R. W., Shipham, M. A., Browell, E. V., Jacob, D. J., and Logan, J. A.: Atmospheric chemistry in the Arctic and Subarctic: influence of natural fires, industrial emissions, and stratospheric inputs, *J. Geophys. Res.*, 97, 16731–16746, 1992.
- Zhang, Q., Streets, D. G., Carmichael, G. R., He, K. B., Huo, H., Kannari, A., Klimont, Z., Park, I. S., Reddy, S., Fu, J. S., Chen, D., Duan, L., Lei, Y., Wang, L. T., and Yao, Z. L.: Asian emissions in 2006 for the NASA INTEX-B mission, *Atmos. Chem. Phys.*, 9, 5131–5153, doi:10.5194/acp-9-5131-2009, 2009.

**Table 1.** Participating models in inter-comparison. Y in the tracers column indicates that idealised tracers from model have been used for the transport and chemistry analysis in Sect. 5, Figures 14-17. '(N.B. MATCH provided the 25-day tracers for inclusion in Figure 13, however, they were excluded from the analysis in Figure 14-17 as the regional model OH field could not be used in the global TOMCAT model to calculate the OH tracers needed to perform the statistical analysis.)

Model	Resolution	Meteorology	Chemistry	Tracers	Reference
CAM4-Chem	1.9° × 2.5°, L56	GEOS-5	MOZART-4	Y	Lamarque et al. (2012)
CAM5-Chem	1.9° × 2.5°, L56	GEOS-5	MOZART-4	Y	Tilmes et al. (2014)
C-IFS	T159, L60	ECMWF-Op.	CB05: 55 species, 85 gas-phase reactions Stratosphere: Operational ECMWF O <sub>3</sub> analysis	Y	Flemming et al. (2014)
GEOS-Chem	2.0° × 2.5°, L47	GEOS-5	~ 100 species HO <sub>2</sub> gamma for uptake by aerosol set to 1	N	Mao et al. (2010) Parrella et al. (2012)
GMI	2.0° × 2.5°, L72	GEOS-5	based on GEOS-Chem	Y	Strahan et al. (2007) Duncan et al. (2007)
LMDZ-INCA	3.75° × 1.8°, L39	ECMWF-Int	85 tracers, 264 gas-phase reactions Stratosphere: Ozonesonde climatology	Y	Hourdin et al. (2006) Szopa et al. (2013)
SMHI MATCH	0.75° × 0.75°, L35	ECMWF-Int.	63 tracers, 110 gas-phase reactions Stratosphere: Monthly means from EU-MACC project (MOZART-4)	N	Andersson et al. (2007) Robertson et al. (1999)
MOZART-4	1.9° × 2.5°, L56	GEOS-5	103 tracers, 196 gas-phase reactions, Stratosphere: O <sub>3</sub> constrained by sondes/satellite	Y	Emmons et al. (2010)
TM5	2.0° × 3.0°, L60	ECMWF-Int.	Modified CB05 scheme: 54 tracers, 104 gas-phase reactions. Stratosphere: O <sub>3</sub> columns nudged to observations	Y	Williams et al. (2013) Huijnen et al. (2010)
TOMCAT	2.8° × 2.8°, L31	ECMWF-Int.	82 Tracers, 229 gas-phase reactions. Stratosphere: 2-D model used for boundary conditions	Y	Monks (2011) Chipperfield (2006)
WRF-Chem	100 and 50 km, L65	WRF	MOZART-4 simulations used as boundary conditions	N	Grell et al. (2005) Fast et al. (2006)

**Table 2.** Measurements used to evaluate the POLMIP models.

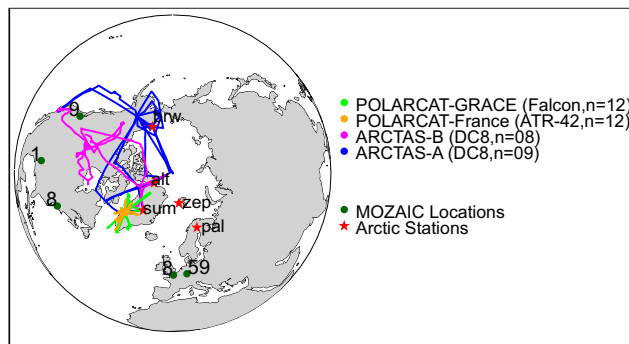
Project	Platform	Period	Location	Species	Method	Uncertainty	Model output frequency	Reference
MOPITT	Satellite	2008	Global	CO	V.6, Near IR & Thermal IR	±9 %	MM	(Deeter et al., 2014)
NOAA ESRLIGMD	Surface	2008 2008	See Fig. 1	CO	Gas chromatography (Flask)	±1 %	MM	(Novelli et al., 1998)
				O <sub>3</sub>	UV absorption	±5 %	MM	(Oltmans and Levy, 1994)
MOZAIC	Aircraft	18/06/08– 14/07/08	Ascent & descent at: Frankfurt, London, Philadelphia, Portland, Dallas	CO	IR absorption	±5 ppbv	HI	Nédélec et al. (2003)
				O <sub>3</sub>	UV absorption	±2 ppbv	HI	Thouret et al. (1998)
POLARCAT-France	Aircraft	30/06/08– 14/07/08	Based in Kangerlussuaq, Greenland	CO	IR absorption	±5 ppbv	HI	Nédélec et al. (2003)
				O <sub>3</sub>	UV absorption	±1 %	HI	Ancellet et al. (2009)
POLARCAT-GRACE	Aircraft	02/07/08– 14/07/08	Based in Kangerlussuaq, Greenland	CO	Vacuum UV fluorescence	±5 %	HI	Gerbig et al. (1999)
				O <sub>3</sub>	UV absorption	±5 %	HI	Roiger et al. (2011)
ARCTAS-A	Aircraft	01/04/08– 19/04/08	Based in Fairbanks, Alaska	CO	TDLIF	±2 %	HI	Sachse and Hill (1987)
				O <sub>3</sub>	Chemiluminescence	±5 %	HI	Weinheimer et al., (1994)
				OH	Laser induced fluorescence	±40 %	HI	Brune et al. (1999)
ARCTAS-B	Aircraft	29/06/08 10/07/08	Based in Cold Lake, Canada	CO	TDLIF	±2 %	HI	Sachse and Hill (1987)
				O <sub>3</sub>	Chemiluminescence	±5 %	HI	Weinheimer et al., (1994)
				OH	Laser Induced Fluorescence	±40 %	HI	Brune et al. (1999)

a IR = infrared

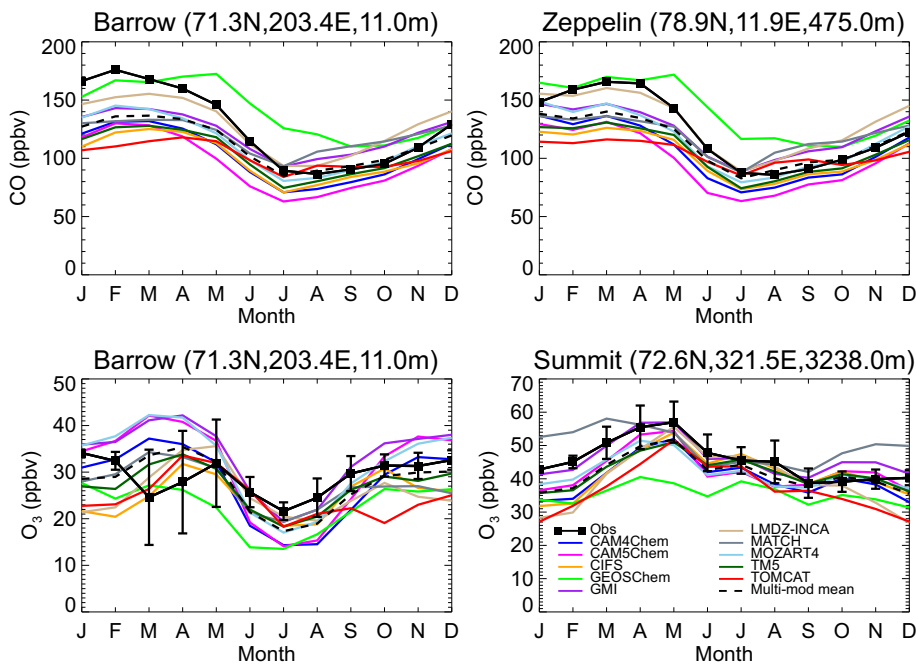
b UV = ultraviolet

c TDLIF = Thermal-Dissociation Laser Induced Fluorescence

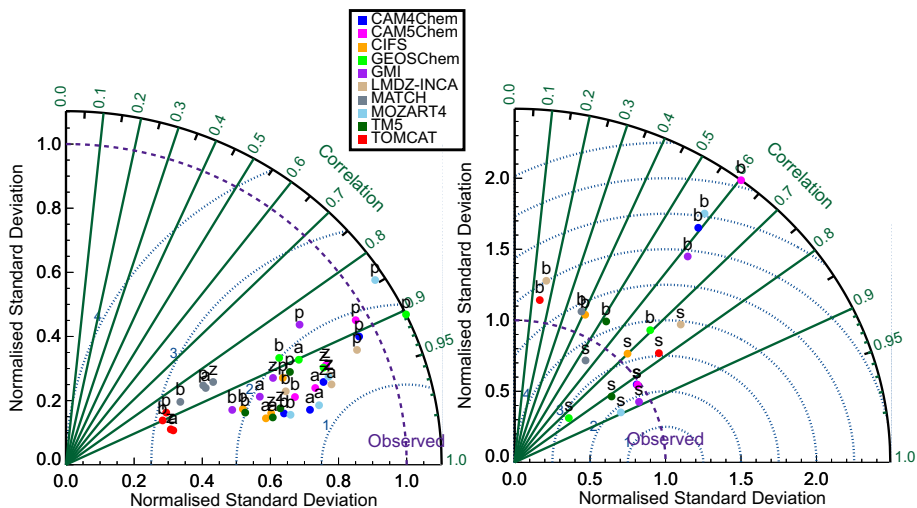
d MM = monthly mean, HI = Hourly instantaneous.



**Figure 1.** Location of surface stations used for comparisons in Figs. 2–3 (alt = Alert, brw = Barrow, pal = Pallas, sum = Summit, zep = Zeppelin), flight tracks of POLARCAT flights used to calculate vertical profiles in Figs. 7–8 (with n = no. of flights used) and locations of the airports where MOZAIC descent and ascent profiles were used over Europe and North America in Fig. 9 (the values denote the number of profiles at each airport).

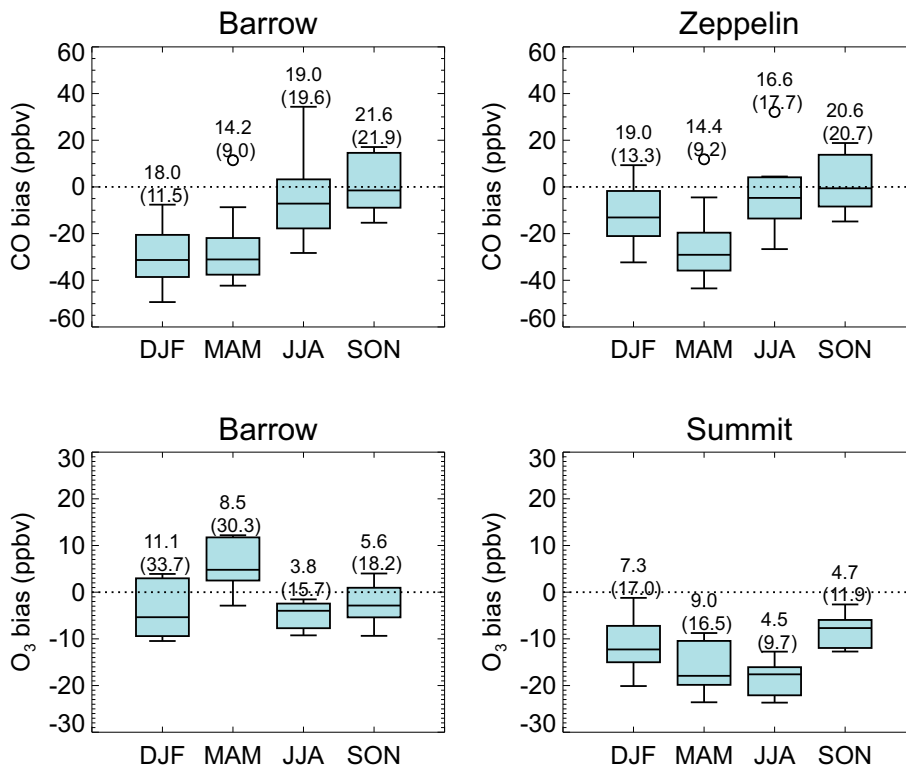


**Figure 2.** Monthly mean 2008 simulated and observed Arctic surface CO (top) at Barrow and Zeppelin and surface O<sub>3</sub> (bottom) at Barrow and Summit. Error bars show 1σ of the daily observed O<sub>3</sub> data.

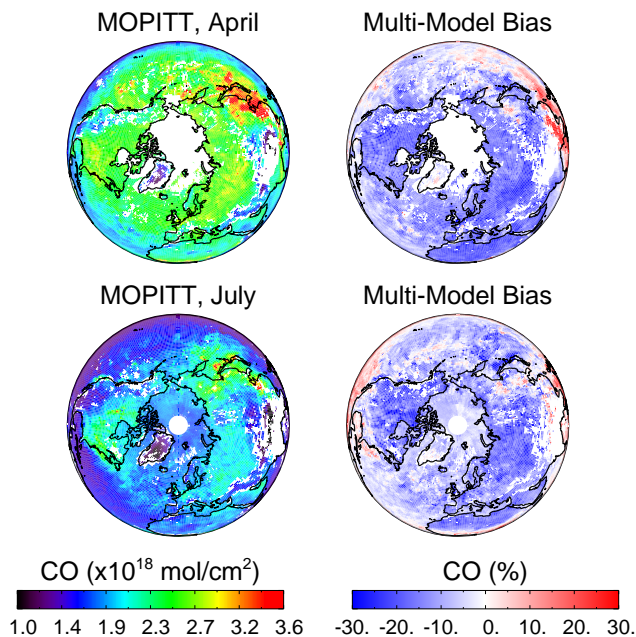


**Figure 3.** Taylor diagrams for Arctic surface comparisons of monthly mean time series of CO (left) and O<sub>3</sub> (right). The letters represent the first letter of the station codes shown in Fig. 1 (a=Alert, b=Barrow, p=Pallas, s=Summit, z=Zeppelin). (Taylor diagrams show the extent to which models capture the observed variability (by the normalised standard deviation (NSD), shown by purple dashed contour), month-to-month changes in concentration (by the correlation ( $r$ ), shown by green solid lines) and the mean model bias as a function of the variations around the annual mean (by the centered root-mean-square difference ( $\text{RMSD} = \frac{1}{n} \sum_{n=1,12} (m_n - \bar{m}) - (o_n - \bar{o})$ ), shown by the blue dotted contours). Perfect agreement between a model and observations would result in a NSD of 1, a correlation of 1 and a RMSD of 0, which is indicated by “Observed” on the Taylor diagram.)

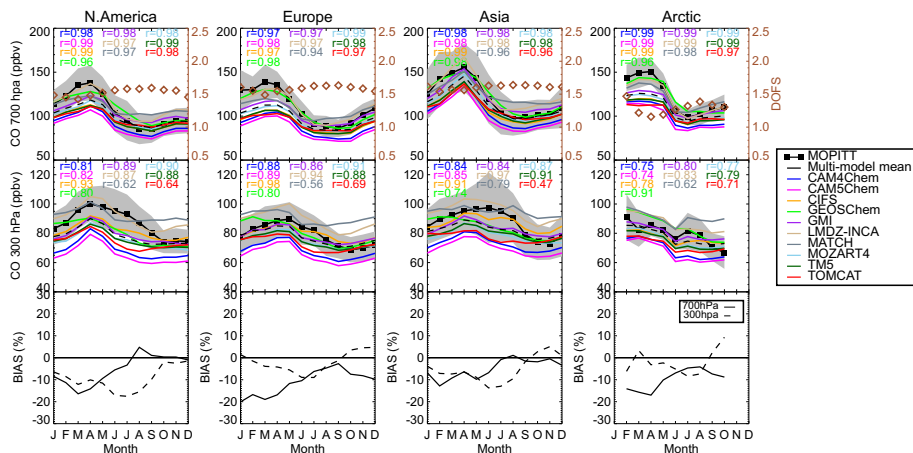




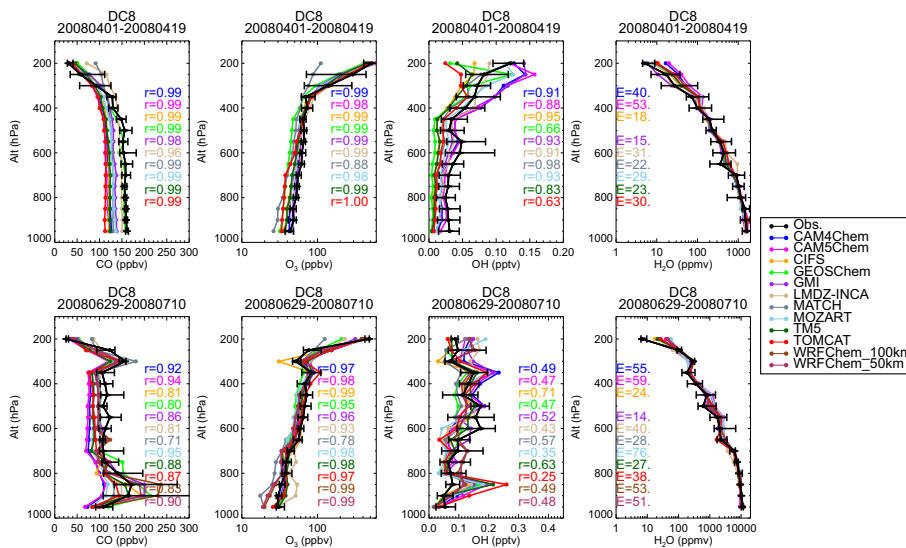
**Figure 4.** Box and whisker plots showing model seasonal biases (ppbv) for CO at Barrow and Zeppelin (top) and O<sub>3</sub> at Barrow and Summit (bottom). The box and whisker plots show the minimum, 25th percentile, median, 75th percentile and maximum values of the sample. The numbers on the plot represent the interquartile range of each sample (IQR = 75th–25th percentiles) in absolute concentrations and as a percent of observed concentrations (in brackets). Outliers which are more than  $1.5 \times$  IQR from the 25th or 75th percentiles are shown by circles.



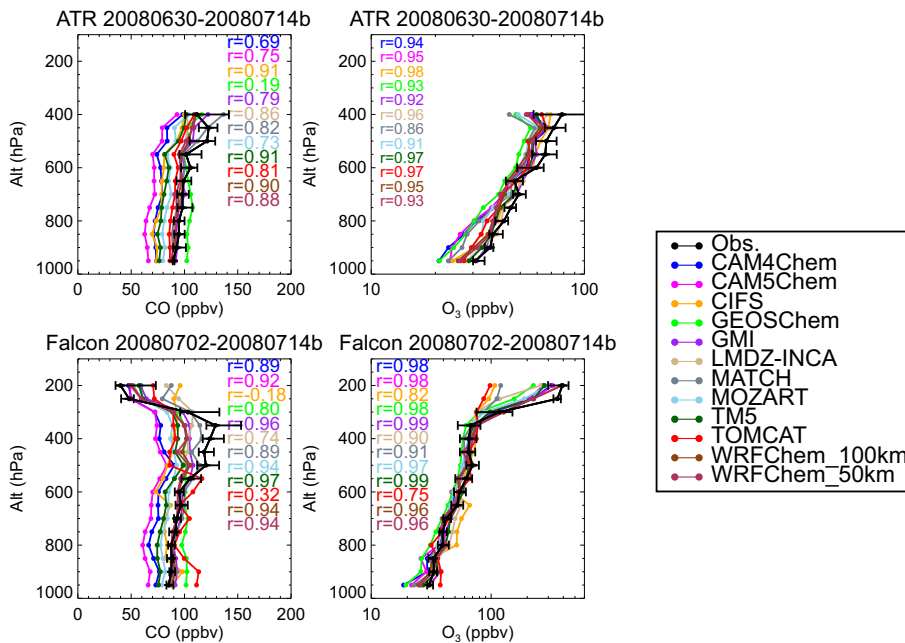
**Figure 5.** MOPITT total tropospheric CO column (left) and multi-model mean percent bias (right) for April (top) and July (bottom) 2008. The models have had the MOPITT averaging kernels applied and retrievals with DOFS less than 1 removed. The multi-model mean CO is calculated from all 10 global models and the MATCH hemispheric model.



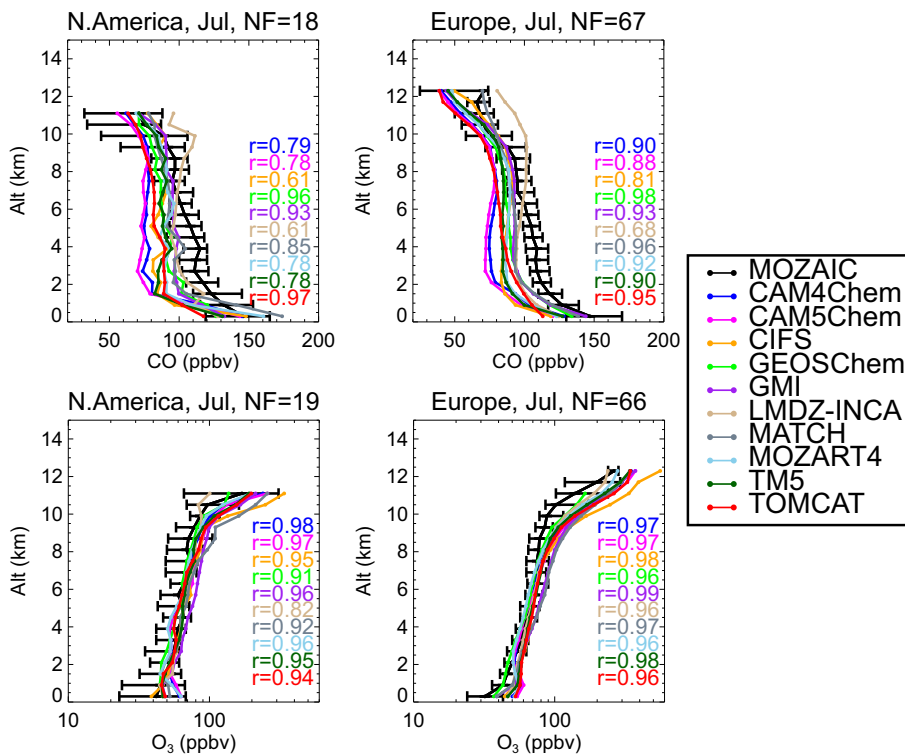
**Figure 6.** Monthly mean 2008 MOPITT retrieved CO compared to models at 700 hPa (top) and 300 hPa (middle). The monthly mean percent bias (bottom) at 700 hPa (solid lines) and 300 hPa (dashed lines) are also shown. The models have had the MOPITT averaging kernels applied and retrievals with DOFS less than 1 removed. The grey shaded area shows 1 $\sigma$  of the MOPITT concentrations over the regions at the relevant pressure level.



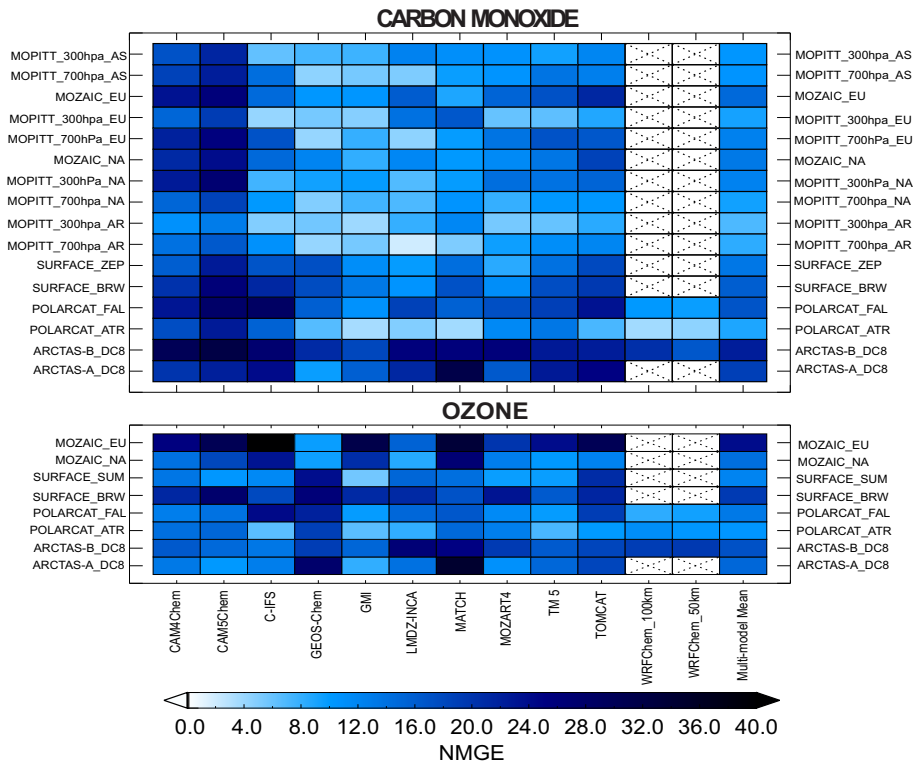
**Figure 7.** DC8 vertical profiles from the spring ARCTAS-A (top) and summer ARCTAS-B (bottom) campaigns in 2008. Median concentrations of CO (left), O<sub>3</sub> (middle), OH (middle) and H<sub>2</sub>O (right), with error bars showing the 25th and 75th percentiles of the observations. Pearson's correlation coefficients are shown for CO, O<sub>3</sub> and OH. NMGE (%) is shown for H<sub>2</sub>O. (N.B. Due to missing data in the GEOS-Chem hourly files the median concentrations in the 200, 250 and 300 hPa bins have been calculated over 463 data points instead of 678 data points which are used for the other models and observations.)



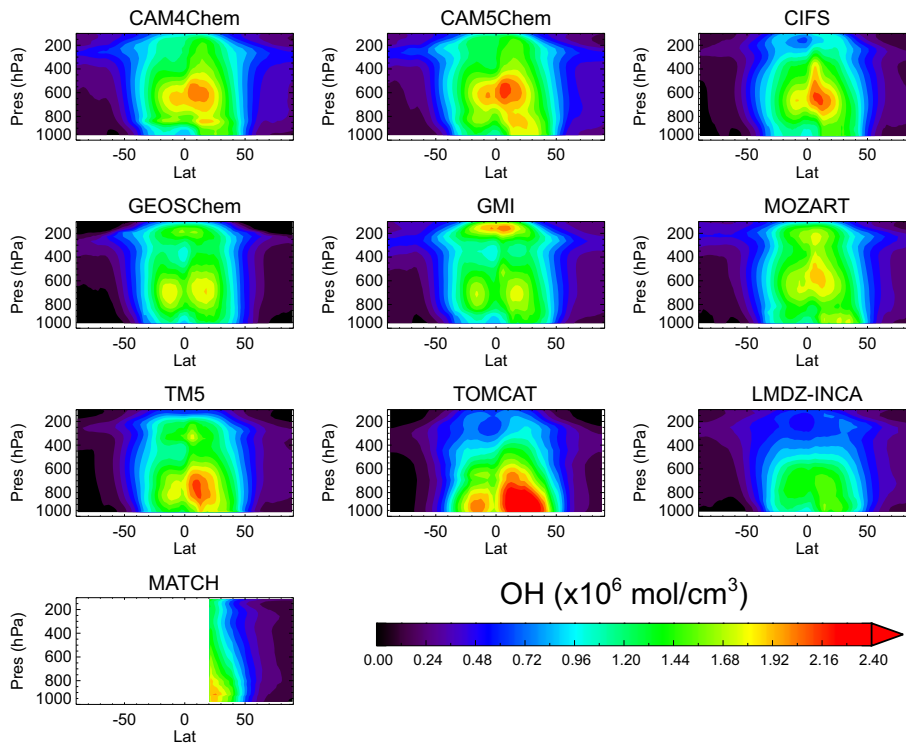
**Figure 8.** Vertical median profiles from the ATR-42 during the summer POLARCAT-France campaign (top) and the Falcon during the summer POLARCAT-GRACE campaign (bottom) for CO (left) and O<sub>3</sub> (right). Pearson's correlation coefficients are also shown. Error bars show the 25th and 75th percentiles of the observations.



**Figure 9.** Vertical median profiles from the MOZIAIC aircraft project made between 18 June–14 July 2008, during ascent and descent in the regions of North America (left) and Europe (right). Observed CO (top) and O<sub>3</sub> (bottom) are compared to hourly simulated concentrations interpolated to MOZIAIC airport location. All data is put into 50 m bins and the error bars show the 25th and 75th percentiles of the observations in each bin. The number of profiles in each location used to calculate the regional profile averages are shown in Fig. 1.

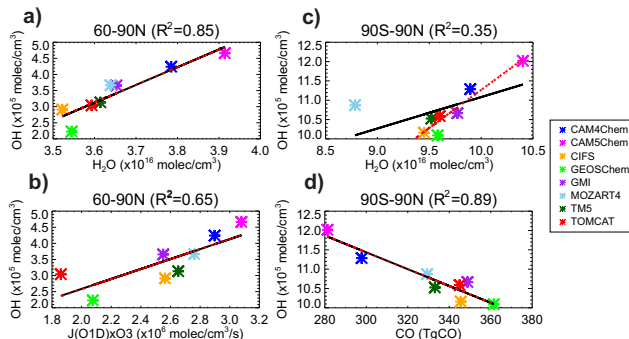


**Figure 10.** Summary of the Normalised Mean Gross Error ( $NMGE = \frac{1}{N} \sum \left( \frac{|mod-obs|}{obs} \right) \times 100$ ) for all models against observations. The errors are shown for MOPITT regionally averaged satellite retrievals at 300 and 700 hPa (where EU = Europe, NA = North America, AS = Asia, AR = Arctic), MOZAIC vertical profiles over source regions (where EU = Europe, NA = North America), at surface sites (where BRW = Barrow, ZEP = Zeppelin, SUM = Summit) and POLARCAT aircraft observations (where POLARCAT\_FAL = POLARCAT-GRACE, POLARCAT\_ATR = POLARCAT-France).

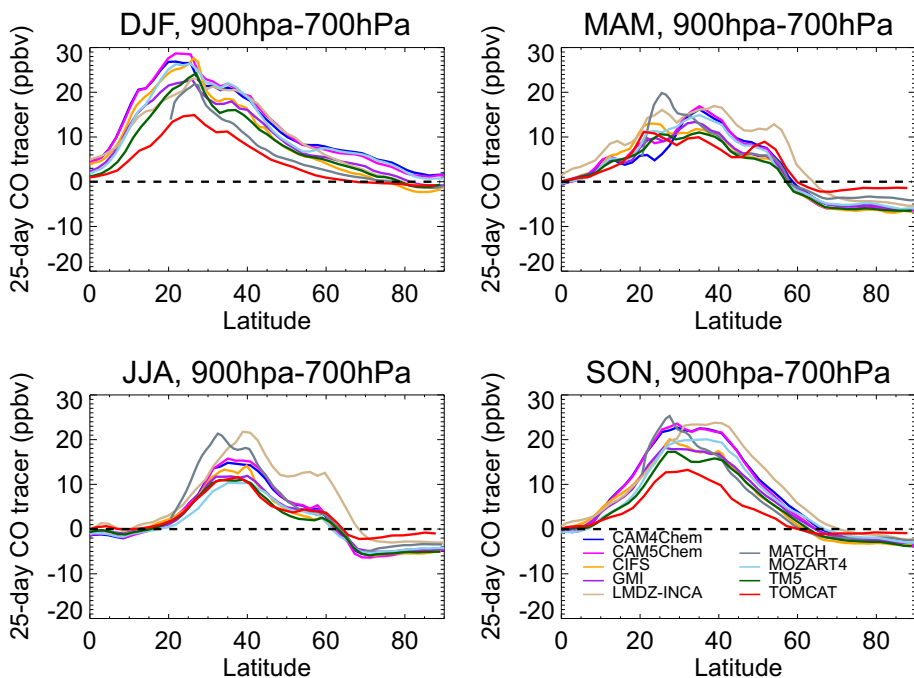


**Figure 11.** Annual zonal mean OH concentrations ( $\times 10^5$  molecules cm<sup>-3</sup>) for 2008 from the POLMIP models.

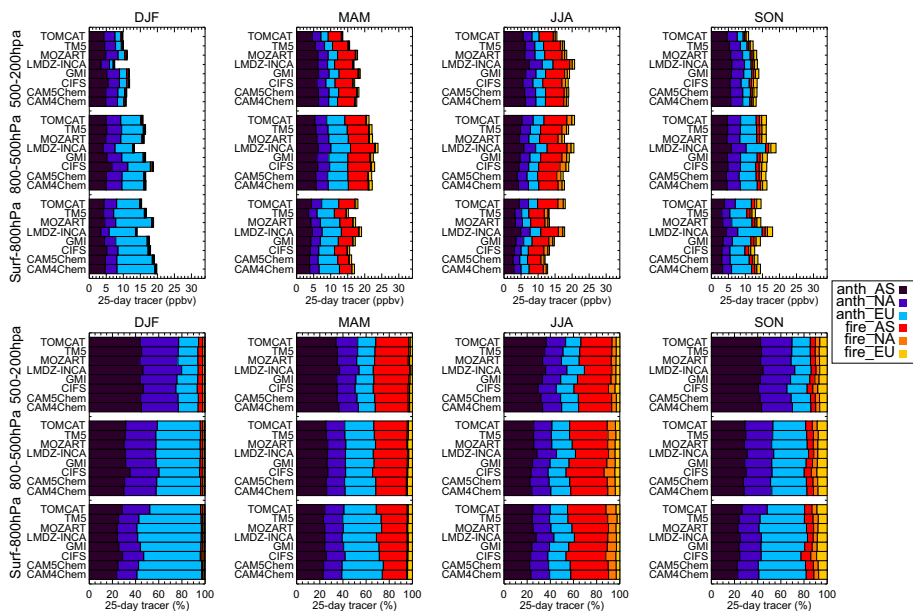




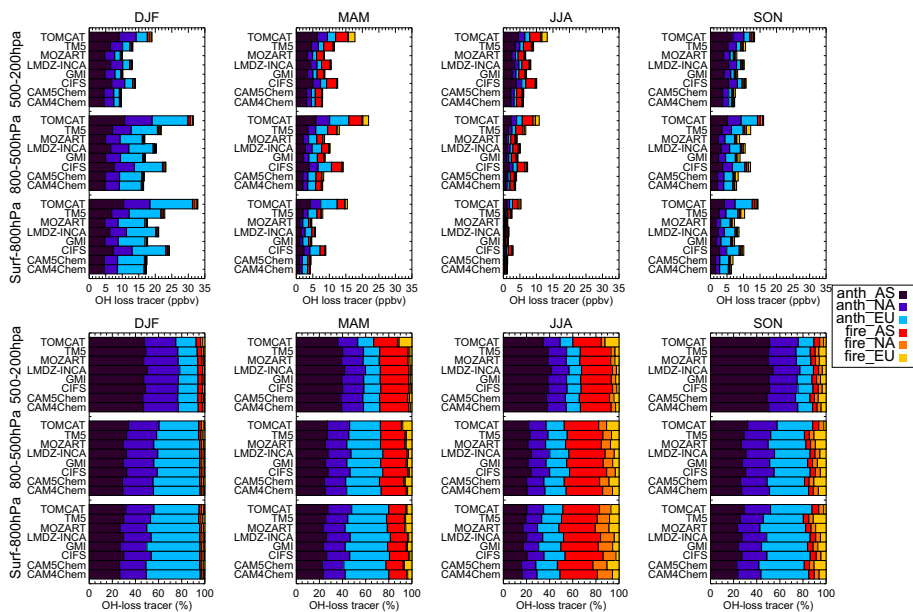
**Figure 12.** Annual mass-weighted tropospheric mean OH concentration (a) against annual mass-weighted mean water vapour in the Arctic, (b) against annual mass-weighted mean photolysis rates of  $J(O^1D)$  multiplied by  $O_3$  concentrations in the Arctic (c) against annual mass-weighted mean water vapour concentrations for the whole troposphere and (d) against the annual mean CO burden for the whole troposphere. The ordinary least squares (OLS) and the iteratively re-weighted least squares (IRLS) regression lines are shown in black and red, respectively. Correlations equal to or above 0.62 are significant at the 90% confidence level (CL) or above. The troposphere has been selected using the 150 ppbv  $O_3$  contour. See Table S1 for more details.



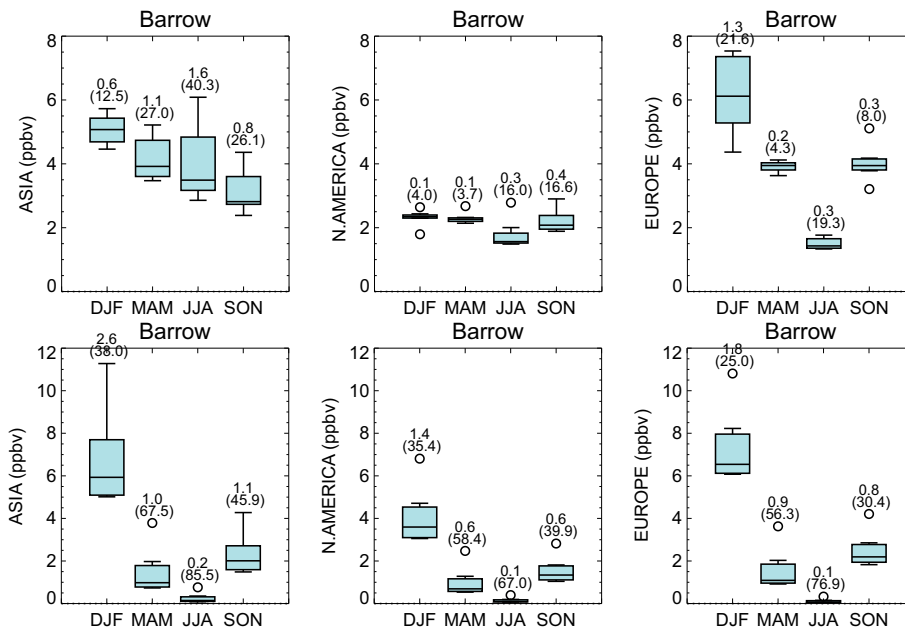
**Figure 13.** Seasonal zonal mean differences in simulated total 25-day fixed-loss tracer (all six anthropogenic and fire tracers have been summed) between 900 and 700 hPa levels. (GEOS-Chem is excluded as the 25-day fixed lifetime tracers experiments were not performed.)



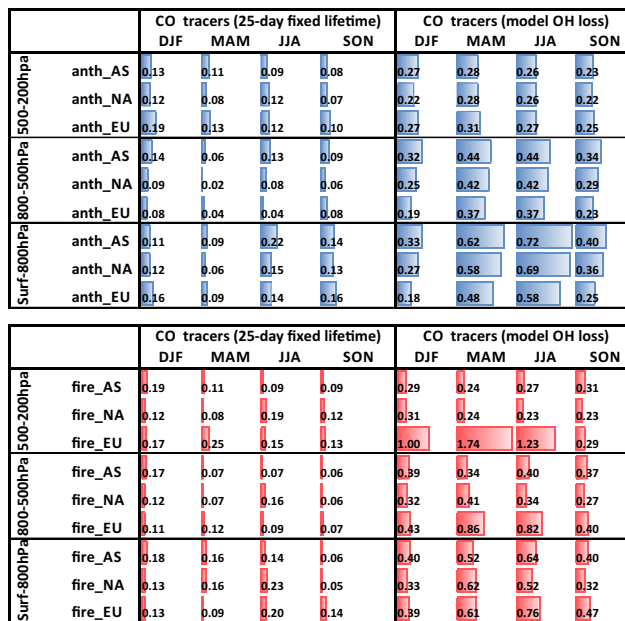
**Figure 14.** Seasonal mean 2008 regional 25-day fixed-loss anthropogenic and biomass burning regional tracers averaged over 3 different altitude bands at latitudes north of 66° N. Contributions shown as absolute concentrations (top) and as a percent of the total CO tracer (bottom). (AS = Asian tracer, NA = North American tracer, EU = European tracer).



**Figure 15.** Seasonal mean 2008 regional OH-loss anthropogenic and biomass burning regional tracers averaged over 3 different altitude bands at latitudes north of 66° N. Contributions shown as absolute concentrations (top) and as a percent of the total CO tracer (bottom). (AS = Asian tracer, NA = North American tracer, EU = European tracer).



**Figure 16.** Box and Whisker plots of the anthropogenic 25-day fixed-loss tracers (top) and the OH-loss tracers (bottom) at Barrow from 8 POLMIP models. The box and whisker plots show the minimum, 25th percentile, median, 75th percentile and maximum values of the sample. Outliers which are more than  $1.5 \times \text{IQR}$  from the 25th or 75th percentiles are shown by circles. The numbers on the plot represent the interquartile range as absolute concentrations and as a percent of the multi-model mean (in brackets).



**Figure 17.** Coefficients of variations ( $\frac{\sigma}{\mu}$ ) calculated from the POLMIP model data binned into altitude bands at latitudes north of  $66^\circ$  N shown in Figs. 14 and 15. The size of the data bars represent the value of the coefficients with anthropogenic coefficients in blue and fire coefficients in red.